

# Descriptive complexity for pictures languages (Extended abstract)

Etienne Grandjean<sup>a</sup>, Frédéric Olive<sup>b</sup>, Gaétan Richard<sup>a</sup>

<sup>a</sup>Université de Caen / ENSICAEN / CNRS - GREYC - Caen, France

<sup>b</sup>Aix-Marseille Université - LIF - Marseille, France

---

## Abstract

This paper deals with descriptive complexity of picture languages of any dimension by syntactical fragments of existential second-order logic.

- We uniformly generalize to any dimension the characterization by Giammarresi et al. [GRST96] of the class of *recognizable* picture languages in existential monadic second-order logic.
- We state several logical characterizations of the class of picture languages recognized in linear time on nondeterministic cellular automata of any dimension. They are the first machine-independent characterizations of complexity classes of cellular automata.

Our characterizations are essentially deduced from normalization results we prove for first-order and existential second-order logics over pictures. They are obtained in a general and uniform framework that allows to extend them to other "regular" structures. Finally, we describe some hierarchy results that show the optimality of our logical characterizations and delineate their limits.

**Keywords:** Picture languages, locality, recognizability, linear time, cellular automata, logical characterizations, existential second-order logic.

---

## Introduction

One goal of descriptive complexity is to establish logical characterizations of natural classes of problems in finite model theory. Many results in this area involve second-order logic (SO) and its restrictions, monadic second-order logic (MSO) and existential second-order logic (ESO). Indeed, there are two lines of research that roughly correspond to either of these restrictions:

(a) *The formal language current.* It starts from the pioneering result by Büchi, Elgot and Trahtenbrot [Büc60], which states that the class of regular languages equals the class of MSO-definable languages; in short, REG = MSO. This line of research aims at characterizing in logic the natural classes of algebraically defined languages (sets of words) or sets of structures (trees, graphs, etc.) defined by finite state recognizability or local properties such as tilings.

---

*Email addresses:* etienne.grandjean@unicaen.fr (Etienne Grandjean), frederic.olive@lif.univ-mrs.fr (Frédéric Olive), gaetan.richard@info.unicaen.fr (Gaétan Richard)

(b) *The computational complexity current.* It originates from another famous result, Fagin’s Theorem [Fag74], which characterizes the class NP as the class of problems definable in ESO.

For many years, both directions of research have produced plenty of results: see *e.g.* [EF95, Lib04] for descriptive complexity of formal languages and [EF95, Imm99, GKL<sup>+</sup>07, Lib04] for the one of complexity classes. However, and this may be surprising, only few connections are known between those two areas of descriptive complexity. Of course, an explanation is that formal language theory has its own purposes that have little to do with complexity theory. In our opinion, the main reason is that while MSO logic *exactly fits* the fundamental notion of *recognizability*, as exemplified in the work of Courcelle [CE12], this logic seems *transversal* to computational complexity. We argue this is due to the intrinsic *locality* that MSO logic *inherits* from first-order logic [Han65, Gai82]. Typically, whereas MSO, or even existential MSO (EMSO), expresses some NP-complete graph problems such as 3-colourability, it cannot express some other ones such as Hamiltonicity (see [Tur84, dR87, Lib04]) or even some tractable graph properties, such as the existence of a perfect matching. In contrast, the situation is very clear on trees as on words: MSO only captures the class of "easiest" problems; an extension of Büchi’s Theorem [TW68] states that a tree language is MSO definable iff it is recognizable by a finite tree automaton

Thus, items (a) and (b) above seem quite separate for problems on words, trees or graphs, in one case (for words and trees) because MSO only expresses easy problems (regular languages), in the other case because MSO and EMSO do *not* correspond to any complexity class over graphs.

What about *picture languages*, that is sets of  $d$ -pictures, i.e.,  $d$ -dimensional words (or coloured grids)? First, notice the following results:

1. In a series of papers culminating in [GRST96], Giammarresi and al. have proved that a 2-picture language is *recognizable*, i.e. is the projection of a local 2-picture language, iff it is definable in EMSO. In short:  $\text{REC}^2 = \text{EMSO}$ .
2. In fact, the class  $\text{REC}^2$  contains some NP-complete problems. More generally, one observes that for each dimension  $d \geq 1$ ,  $\text{REC}^d$  can be defined as the class of  $d$ -picture languages recognized by nondeterministic  $d$ -dimensional *cellular automata* in *constant time*<sup>1</sup>.

In some sense, the present paper is an attempt to bridge the gap between the formal language current of descriptive complexity involving MSO and the computational complexity current that involves ESO. This paper originates from two questions about word/picture languages:

1. How can we generalize the proof of the above-mentioned theorem of Giammarresi and al. to any dimension? That is, can we establish the equality  $\text{REC}^d = \text{EMSO}$  for any  $d \geq 1$ ?
2. Can we obtain logical characterizations of time complexity classes of cellular automata<sup>2</sup>?

The paper addresses both questions; it also compares, in a common framework, the point of view of formal language theory with that of computational complexity. A  $d$ -picture language over an alphabet  $\Sigma$  is a set

---

<sup>1</sup>That means: for such a picture language  $L$ , there is some constant integer  $c$  such that each computation stops at instant  $c$ , and  $p \in L$  iff it has at least one computation whose final configuration is accepting in the following sense: *all the cells are in an accepting state*.

<sup>2</sup>This originates from a question that J. Mazoyer asked us in 2000 (personal communication): give a logical characterization of the linear time complexity class of nondeterministic cellular automata.

of  $d$ -pictures  $p : [1, n]^d \rightarrow \Sigma$ , i.e.,  $d$ -dimensional  $\Sigma$ -words<sup>3</sup>. There are two natural manners to represent a  $d$ -picture  $p$  as a first-order structure:

- as a *pixel structure*: on the *pixel* domain  $[1, n]^d$ ;
- as a *coordinate structure*: on the *coordinate* domain  $[1, n]$ .

Significantly, these two representations respectively correspond to the two above-mentioned points of view as shown by our results.

*Our results:*

We establish two kinds of logical characterizations of  $d$ -picture languages, for all dimensions  $d \geq 1$ :

1. *On pixel structures*:  $\text{REC}^d = \text{ESO}(\text{arity } 1) = \text{ESO}(\text{var } 1) = \text{ESO}(\forall^1, \text{arity } 1)$ . That means a  $d$ -picture language is *recognizable* iff it is definable in monadic ESO (resp. in ESO with 1 first-order variable, or in monadic ESO with 1 universally quantified first-order variable).
2. *On coordinate structures*:  $\text{NLIN}_{\text{ca}}^d = \text{ESO}(\text{var } d + 1) = \text{ESO}(\forall^{d+1}, \text{arity } d + 1)$ ; that means a  $d$ -picture language is recognized by a nondeterministic  $d$ -dimensional cellular automaton in *linear time* iff it is definable in ESO with  $d + 1$  distinct first-order variables (resp. ESO with second-order variables of arity at most  $d + 1$  and a prenex first-order part of prefix  $\forall^{d+1}$ ).

Both items (1) and (2) are easy consequences of *normalization results* of, respectively, first-order and ESO logics we prove over picture languages. In particular, the "normalization" equality  $\text{ESO}(\text{arity } 1) = \text{ESO}(\forall^1, \text{arity } 1)$  is a consequence of the fact that on pixel structures (and more generally, on structures that consist of bijective functions and unary relations) any first-order formula is equivalent to a boolean combination of "cardinality" formulas of the form: "there exists  $k$  distinct elements  $x$  such that  $\psi(x)$ ", where  $\psi$  is a quantifier-free formula with only *one* variable. The "normalization" equality explicitly expresses the "local" feature of MSO on pictures and can be generalized to other "regular" structures.

Thirdly, in contrast with (1) and (2), we establish several *strict hierarchy* results for any fixed  $d \geq 2$  and for  $d$ -picture languages represented by coordinate structures: in particular, we prove

$$\text{ESO}(\text{var } d - 1) \subsetneq \text{REC}^d \subsetneq \text{ESO}(\text{var } d) = \text{NLIN}_{\text{ca}}^d$$

and  $\text{ESO}(\text{var } d) \subsetneq \text{ESO}(\text{arity } d)$ . Here, the fact that *no* natural restriction of ESO (in particular,  $\text{ESO}(\text{var } d)$ ) exactly captures the class  $\text{REC}^d$  for the coordinate representation of pictures seems to us the symptom of the large expressiveness of this logic: significantly, the *strict inclusion*  $\text{ESO}(\text{var } d) \subsetneq \text{ESO}(\text{arity } d)$  for  $d$ -pictures in the coordinate representation, if  $d \geq 2$ , strongly contrasts with the *equality*  $\text{ESO}(\text{arity } 1) = \text{ESO}(\text{var } 1)$  that holds in the pixel representation.

In this document, we sketch many proofs and omit other ones, in particular, the proofs of hierarchy theorems, in Section 4. However, the very technical proof of the normalization of the logic  $\text{ESO}(\forall d, \text{arity } d)$  on coordinate encodings of  $(d - 1)$ -pictures (for  $d \geq 2$ ) is completely described in the appendix.

---

<sup>3</sup>More generally, the domain of a  $d$ -picture is of the "rectangular" form  $[1, n_1] \times \dots \times [1, n_d]$ . For simplicity and uniformity of presentation, we have chosen to present the results of this paper in the particular case of "square" pictures of domain  $[n]^d$ . Fortunately, our results also hold with the same proofs for general domains  $[1, n_1] \times \dots \times [1, n_d]$ .

## 1. Preliminaries

In the definitions below and all along the paper, we denote by  $\Sigma, \Gamma$  some finite alphabets and by  $d$  a positive integer. For any positive integer  $n$ , we set  $[n] := \{1, \dots, n\}$ . We are interested in sets of pictures of any fixed dimension  $d$ .

**Definition 1.1.** A  $d$ -dimensional picture or  $d$ -picture on  $\Sigma$  is a function  $p : [n]^d \rightarrow \Sigma$  where  $n$  is a positive integer. The set  $\text{dom}(p) = [n]^d$  is called the **domain** of picture  $p$  and its elements are called **points**, **pixels** or **cells** of the picture. A set of  $d$ -pictures on  $\Sigma$  is called a  $d$ -dimensional language, or  $d$ -language, on  $\Sigma$ .

Notice that 1-pictures on  $\Sigma$  are nothing but nonempty words on  $\Sigma$ .

### 1.1. Pictures as model theoretic structures

Along the paper, we will often describe  $d$ -languages as sets of models of logical formulas. To allow this point of view, we must settle on an encoding of  $d$ -pictures as model theoretic structures.

For logical aspects of this paper, we refer to the usual definitions and notations in logic and finite model theory (see [EF95] or [Lib04], for instance). A **signature** (or **vocabulary**)  $\sigma$  is a finite set of relation and function symbols each of which has a fixed arity. A (finite) **structure**  $S$  of vocabulary  $\sigma$ , or  $\sigma$ -structure, consists of a finite domain  $D$  of cardinality  $n \geq 1$ , and, for any symbol  $s \in \sigma$ , an interpretation of  $s$  over  $D$ , often denoted by  $s$  for simplicity. The tuple of the interpretations of the  $\sigma$ -symbols over  $D$  is called the **interpretation** of  $\sigma$  over  $D$  and, when no confusion results, it is also denoted  $\sigma$ . The **cardinality of a structure** is the cardinality of its domain. For any signature  $\sigma$ , we denote by  $\text{STRUC}(\sigma)$  the class of (finite)  $\sigma$ -structures. We write  $\text{MODELS}(\Phi)$  the set of  $\sigma$ -structures which satisfy some fixed formula  $\Phi$ . We will often deal with **tuples** of objects. We denote them by bold letters.

There are two natural manners to represent a picture by some logical structure: on the domain of its pixels, or on the domain of its coordinates. This gives rise to the following definitions:

**Definition 1.2.** Given  $p : [n]^d \rightarrow \Sigma$ , we denote by  $\text{pixel}^d(p)$  the structure

$$\text{pixel}^d(p) = ([n]^d, (Q_s)_{s \in \Sigma}, (\text{succ}_i)_{i \in [d]}, (\text{min}_i)_{i \in [d]}, (\text{max}_i)_{i \in [d]}).$$

Here:

- $\text{succ}_j$  is the **(cyclic) successor function** according to the  $j^{\text{th}}$  dimension of  $[n]^d$ , that is: for each  $a = (a_i)_{i \in [d]} \in [n]^d$ ,  $\text{succ}_j(a) = a^{(j)}$  where  $a_i^{(j)}$ , the  $i^{\text{th}}$  component of  $a^{(j)}$ , equals  $a_i$ , the  $i^{\text{th}}$  component  $a_i$  of  $a$ , except the  $j^{\text{th}}$  one which equals the cyclic successor of the  $j^{\text{th}}$  component  $a_j$  of  $a$ . More formally:
  - $a_j^{(j)} = a_j + 1$  if  $a_j < n$ , and  $a_j^{(j)} = 1$  otherwise;
  - $a_i^{(j)} = a_i$  for each  $i \neq j$ .
- the  $\text{min}_i$ 's,  $\text{max}_i$ 's and  $Q_s$ 's are the following unary (monadic) relations:
  - $\text{min}_i = \{a \in [n]^d : a_i = 1\};$
  - $\text{max}_i = \{a \in [n]^d : a_i = n_i\};$
  - $Q_s = \{a \in [n]^d : p(a) = s\}.$

**Definition 1.3.** Given  $p : [n]^d \rightarrow \Sigma$ , we denote by  $\text{coord}^d(p)$  the structure

$$\text{coord}^d(p) = \langle [n], (Q_s)_{s \in \Sigma}, <, \text{succ}, \text{min}, \text{max} \rangle. \quad (1)$$

Here:

- Each  $Q_s$  is a  $d$ -ary relation symbol interpreted as the set of cells of  $p$  labelled by an  $s$ . In other words:  $Q_s = \{a \in [n]^d : p(a) = s\}$ .
- $<, \text{min}, \text{max}$  are predefined relation symbols of respective arities 2, 1, 1, that are interpreted, respectively, as the sets  $\{(i, j) : 1 \leq i < j \leq n\}$ ,  $\{1\}$  and  $\{n\}$ .
- $\text{succ}$  is a unary function symbol interpreted as the cyclic successor. (That is:  $\text{succ}(i) = i + 1$  for  $i < n$  and  $\text{succ}(n) = 1$ .)

For a  $d$ -language  $L$ , we set  $\text{pixel}^d(L) = \{\text{pixel}^d(p) : p \in L\}$  and  $\text{coord}^d(L) = \{\text{coord}^d(p) : p \in L\}$ .

### 1.2. Logics under consideration

Let us now come to the logics involved in the paper. All formulas considered hereafter belong to *relational Existential Second-Order logic*. Given a signature  $\sigma$ , indifferently made of relational and functional symbols, a relational existential second-order formula of signature  $\sigma$  has the shape  $\Phi \equiv \exists \mathbf{R} \varphi(\sigma, \mathbf{R})$ , where  $\mathbf{R} = (R_1, \dots, R_k)$  is a tuple of relational symbols and  $\varphi$  is a first-order formula of signature  $\sigma \cup \{\mathbf{R}\}$ . We denote by  $\text{ESO}^\sigma$  the class thus defined. We will often omit to mention  $\sigma$  for considerations on these logics that do not depend on the signature. Hence, **ESO** stands for the class of all formulas belonging to  $\text{ESO}^\sigma$  for some  $\sigma$ .

We will pay great attention to several variants of ESO. In particular, we will distinguish formulas of type  $\Phi \equiv \exists \mathbf{R} \varphi(\sigma, \mathbf{R})$  according to:

- the number of distinct first-order variables involved in  $\varphi$ ,
- the arity of the second-order symbols  $R \in \mathbf{R}$ , and
- the quantifier prefix of some prenex form of  $\varphi$ .

With the logic  $\text{ESO}^\sigma(\forall^d, \text{arity } \ell)$ , we control these three parameters: it is made of formulas of which first-order part is prenex with a universal quantifier prefix of length  $d$ , and where existentially quantified relation symbols are of arity at most  $\ell$ . In other words,  $\text{ESO}^\sigma(\forall^d, \text{arity } \ell)$  collects formulas of shape:

$$\exists \mathbf{R} \forall \mathbf{x} \theta(\sigma, \mathbf{R}, \mathbf{x})$$

where  $\theta$  is quantifier free,  $\mathbf{x}$  is a  $d$ -tuple of first-order variables, and  $\mathbf{R}$  is a tuple of relation symbols of arity smaller than  $\ell$ . Relaxing some constraints of the above definition, we set:

$$\text{ESO}^\sigma(\forall^d) = \bigcup_{\ell > 0} \text{ESO}^\sigma(\forall^d, \text{arity } \ell) \text{ and } \text{ESO}^\sigma(\text{arity } \ell) = \bigcup_{d > 0} \text{ESO}^\sigma(\forall^d, \text{arity } \ell).$$

Finally, we write  $\text{ESO}^\sigma(\text{var } d)$  for the class of formulas that involve at most  $d$  first-order variables, thus focusing on the sole number of distinct first-order variables (possibly quantified several times).

## 2. A logical characterization of REC

In order to define a notion of locality based on sub-pictures we need to mark the border of each picture.

**Definition 2.1.** By  $\Gamma^\#$  we denote the alphabet  $\Gamma \cup \{\#\}$  where  $\#$  is a special symbol not in  $\Gamma$ . Let  $p$  be any  $d$ -picture of domain  $[n]^d$  on  $\Gamma$ . The **bordered  $d$ -picture** of  $p$ , denoted by  $p^\#$ , is the  $d$ -picture  $p^\# : [0, n+1]^d \rightarrow \Gamma^\#$  defined by

$$p^\#(a) = \begin{cases} p(a) & \text{if } a \in \text{dom}(p); \\ \# & \text{otherwise.} \end{cases}$$

Here, "otherwise" means that  $a$  is on the border of  $p^\#$ , that is, some component  $a_i$  of  $a$  is 0 or  $n+1$ .

Let us now define our notion of *local picture language*. It is based on some sets of allowed patterns (called tiles) of the bordered pictures.

- Definition 2.2.**
1. Given a  $d$ -picture  $p$  and an integer  $j \in [d]$ , two cells  $a = (a_i)_{i \in [d]}$  and  $b = (b_i)_{i \in [d]}$  of  $p$  are  **$j$ -adjacent** if they have the same coordinates, except the  $j^{\text{th}}$  one for which  $|a_j - b_j| = 1$ .
  2. A **tile** for a  $d$ -language  $L$  on  $\Gamma$  is a couple in  $(\Gamma^\#)^2$ .
  3. A picture  $p$  is  **$j$ -tiled** by a set of tiles  $\Delta \subseteq (\Gamma^\#)^2$  if for any two  $j$ -adjacent points  $a, b \in \text{dom}(p^\#)$ :  $(p^\#(a), p^\#(b)) \in \Delta$ .
  4. Given  $d$  sets of tiles  $\Delta_1, \dots, \Delta_d \subseteq (\Gamma^\#)^2$ , a  $d$ -picture  $p$  is **tiled by  $(\Delta_1, \dots, \Delta_d)$**  if  $p$  is  $j$ -tiled by  $\Delta_j$  for each  $j \in [d]$ .
  5. We denote by  $L(\Delta_1, \dots, \Delta_d)$  the set of  $d$ -pictures on  $\Gamma$  that are tiled by  $(\Delta_1, \dots, \Delta_d)$ .
  6. A  $d$ -language  $L$  on  $\Gamma$  is **local** if there exist  $\Delta_1, \dots, \Delta_d \subseteq (\Gamma^\#)^2$  such that  $L = L(\Delta_1, \dots, \Delta_d)$ . We then say that  $L$  is  **$(\Delta_1, \dots, \Delta_d)$ -local**, or  **$(\Delta_1, \dots, \Delta_d)$ -tiled**.

**Definition 2.3.** A  $d$ -language  $L$  on  $\Sigma$  is **recognizable** if it is the projection of a local  $d$ -language over an alphabet  $\Gamma$ . It means there exist a surjective function  $\pi : \Gamma \rightarrow \Sigma$  and a local  $d$ -language  $L_{\text{loc}}$  on  $\Gamma$  such that  $L = \{\pi \circ p : p \in L_{\text{loc}}\}$ .

Because of Definition 2.2, it also means there exist a surjective function  $\pi : \Gamma \rightarrow \Sigma$  and  $d$  subsets  $\Delta_1, \dots, \Delta_d$  of  $(\Gamma^\#)^2$  such that

$$L = \{\pi \circ p : p \in L(\Delta_1, \dots, \Delta_d)\}.$$

We write  $\text{REC}^d$  for the class of recognizable  $d$ -languages.

**Remark 2.4.** Our notion of locality is weaker than the one given by Giammarresi and al. [GRST96]. But this doesn't affect the meaning of recognizability, which coincides with that used in [GRST96]. This confirms the robustness of this latter notion.

A characterization of recognizable languages of dimension 2 by existential monadic second-order logic was proved by by Giammarresi et al. [GRST96]. They established:

**Theorem 2.5 ([GRST96]).** For any 2-dimensional language  $L$ ,  $L \in \text{REC}^2 \Leftrightarrow \text{pixel}^2(L) \in \text{ESO}(\text{arity } 1)$ .

In this section, we come back to this result. We simplify its proof, refine the logic it involves, and generalize its scope to any dimension.

**Theorem 2.6.** *For any  $d > 0$  and any  $d$ -language  $L$ , the following assertions are equivalent:*

1.  $L \in \text{REC}^d$ ;
2.  $\text{pixel}^d(L) \in \text{ESO}(\forall^1, \text{arity } 1)$ ;
3.  $\text{pixel}^d(L) \in \text{ESO}(\text{arity } 1)$ .

Proposition 2.7 states the equivalence  $1 \Leftrightarrow 2$ . In Proposition 3.9, we establish the normalization  $\text{ESO}(\text{arity } 1) = \text{ESO}(\forall^1, \text{arity } 1)$  on pixel structures, from which the equivalence  $2 \Leftrightarrow 3$  immediately follows.

**Proposition 2.7.** *For any  $d > 0$  and any  $d$ -language  $L$  on  $\Sigma$ :  $L \in \text{REC}^d \Leftrightarrow \text{pixel}^d(L) \in \text{ESO}(\forall^1, \text{arity } 1)$ .*

*Sketch of Proof.*  $\Rightarrow$  A picture belongs to  $L$  if there exists a tiling of its domain whose projection coincides with its content. In the logic involved in the proposition, the “arity 1” corresponds to formulating the existence of the tiling, while the  $\forall^1$  is the syntactic resource needed to express that the tiling behaves as expected. Let us detail these considerations.

By Definition 2.2, there exist an alphabet  $\Gamma$  (which can be assumed disjoint from  $\Sigma$ ), a surjective function  $\pi : \Gamma \rightarrow \Sigma$  and  $d$  subsets  $\Delta_1, \dots, \Delta_d \subseteq (\Gamma^\#)^2$  such that  $L$  is the set  $\{\pi \circ p' : p' \in L(\Delta_1, \dots, \Delta_d)\}$ .

The belonging of a picture  $p' : [n]^d \rightarrow \Gamma$  to  $L(\Delta_1, \dots, \Delta_d)$  is easily expressed on  $\text{pixel}^d(p') = \langle [n]^d, (Q_s)_{s \in \Gamma}, \dots \rangle$  with a first-order formula which asserts, for each dimension  $i \in [d]$ , that for any pixel  $x$  of  $p'$ , the couple  $(x, \text{succ}_i(x))$  can be tiled with some element of  $\Delta_i$ . Because it deals with each cell  $x$  separately, this formula has the form  $\forall x \Psi(x, (Q_s)_{s \in \Gamma})$ , where  $\Psi$  is quantifier-free.

Now, a picture  $p : [n]^d \rightarrow \Sigma$  belongs to  $L$  iff it results from a  $\pi$ -renaming of a picture  $p' \in L(\Delta_1, \dots, \Delta_d)$ . It means there exists a  $\Gamma$ -labeling of  $p$  (that is, a tuple  $(Q_s)_{s \in \Gamma}$  of subsets of  $[n]^d$ ) corresponding to a picture of  $L(\Delta_1, \dots, \Delta_d)$  (i.e. fulfilling  $\forall x \Psi(x, (Q_s)_{s \in \Gamma})$ ) and from which the actual  $\Sigma$ -labeling of  $p$  (that is, the subsets  $(Q_s)_{s \in \Sigma}$ ) is obtained via  $\pi$  (easily expressed by a formula of the form  $\forall x \Psi'(x, (Q_s)_{s \in \Sigma}, (Q_s)_{s \in \Gamma})$ ).

Finally, the formula  $(\exists (Q_s)_{s \in \Gamma}) \forall x : \Psi \wedge \Psi'$  conveys the desired property and fits the required form.

$\Leftarrow$  In order to prove the converse implication, it is convenient to first normalize the sentences of  $\text{ESO}(\forall^1, \text{arity } 1)$ . This is the role of the technical result below, which asserts that on pixel encodings, each such sentence can be rewritten in a very local form where the first-order part alludes only pairs of adjacent pixels of the bordered picture. We state it without proof:

**Fact 2.8.** *On pixel structures, any  $\varphi \in \text{ESO}(\forall^1, \text{arity } 1)$  is equivalent to a sentence of the form:*

$$\exists \mathbf{U} \forall x \bigwedge_{i \in [d]} \left\{ \begin{array}{lll} \min_i(x) & \rightarrow & m_i(x) \wedge \\ \max_i(x) & \rightarrow & M_i(x) \wedge \\ \neg \max_i(x) & \rightarrow & \Psi_i(x) \end{array} \right\}. \quad (2)$$

Here,  $\mathbf{U}$  is a list of monadic relation variables and  $m_i, M_i, \Psi_i$  are quantifier-free formulas such that

- atoms of  $m_i$  and  $M_i$  have all the form  $Q(x)$ ;
- atoms of  $\Psi_i$  have all the form  $Q(x)$  or  $Q(\text{succ}_i(x))$ ,

where, in both cases,  $Q \in \{(Q_s)_{s \in \Sigma}, \mathbf{U}\}$ .

Now, consider  $L$  such that  $\text{pixel}^d(L) \in \text{ESO}(\forall^1, \text{arity } 1)$ . Fact 2.8 ensures that  $\text{pixel}^d(L)$  is characterized by a sentence of the form (2) above. We have to prove that  $L$  is the projection of some local  $d$ -language  $L_{\text{loc}}$  on some alphabet  $\Gamma$ , that is a  $(\Delta_1, \dots, \Delta_d)$ -tiled language for some  $\Delta_1, \dots, \Delta_d \subseteq \Gamma^2$ . Let  $U_1, \dots, U_k$  denote the list of (distinct) elements of the set  $\{(Q_s)_{s \in \Sigma}, U\}$  of unary relation symbols of  $\varphi$  so that the first ones  $U_1, \dots, U_m$  are the  $Q_s$ 's (here,  $\min_i$  and  $\max_i$  symbols are excluded). The trick is to put each subformula  $m_i(x)$ ,  $M_i(x)$  and  $\Psi_i(x)$  of  $\varphi$  into its *complete disjunctive normal form* with respect to  $U_1, \dots, U_k$ . Typically, each subformula  $\Psi_i(x)$  whose atoms are of the form  $U_j(x)$  or  $U_j(\text{succ}_i(x))$ , for some  $j \in [k]$ , is transformed into the following "complete disjunctive normal form":

$$\bigvee_{(\epsilon, \epsilon') \in \Delta_i} \left( \bigwedge_{j \in [k]} \epsilon_j U_j(x) \wedge \bigwedge_{j \in [k]} \epsilon'_j U_j(\text{succ}_i(x)) \right). \quad (3)$$

Here, the following conventions are adopted:

- $\epsilon = (\epsilon_1, \dots, \epsilon_k) \in \{0, 1\}^k$  and similarly for  $\epsilon'$ ;
- for any atom  $\alpha$  and any bit  $\epsilon_j \in \{0, 1\}$ ,  $\epsilon_j \alpha$  denotes the literal  $\alpha$  if  $\epsilon_j = 1$ , the literal  $\neg \alpha$  otherwise.

For  $\epsilon \in \{0, 1\}^k$ , we denote by  $\Theta_\epsilon(x)$  the "complete conjunction"  $\bigwedge_{j \in [k]} \epsilon_j U_j(x)$ . Intuitively,  $\Theta_\epsilon(x)$  is a complete description of  $x$  and the set  $\Gamma = \bigcup_{i \in [m]} \{0^{i-1} 10^{m-i}\} \times \{0, 1\}^{k-m}$  is the set of possible colors (remember that the  $Q_s$ 's that are the  $U_j$ 's for  $j \in [m]$  form a partition of the domain). The complete disjunctive normal form (3) of  $\Psi_i(x)$  can be written into the suggestive form

$$\bigvee_{(\epsilon, \epsilon') \in \Delta_i} (\Theta_\epsilon(x) \wedge \Theta_{\epsilon'}(\text{succ}_i(x))).$$

If each subformula  $m_i(x)$  and  $M_i(x)$  of  $\varphi$  is similarly put into complete disjunctive normal form, that is  $\bigvee_{(\sharp, \epsilon) \in \Delta_i} \Theta_\epsilon(x)$  and  $\bigvee_{(\epsilon, \sharp) \in \Delta_i} \Theta_\epsilon(x)$ , respectively (there is no ambiguity in our implicit definition of the  $\Delta_i$ 's, since  $\sharp \notin \Gamma$ ), then the above sentence (2) equivalent to  $\varphi$  becomes the following equivalent sentence:

$$\varphi' = \exists U \forall x \bigwedge_{i \in [d]} \left\{ \begin{array}{ll} \min_i(x) & \rightarrow \bigvee_{(\sharp, \epsilon) \in \Delta_i} \Theta_\epsilon(x) \\ \max_i(x) & \rightarrow \bigvee_{(\epsilon, \sharp) \in \Delta_i} \Theta_\epsilon(x) \\ \neg \max_i(x) & \rightarrow \bigvee_{(\epsilon, \epsilon') \in \Delta_i} (\Theta_\epsilon(x) \wedge \Theta_{\epsilon'}(\text{succ}_i(x))) \end{array} \right\} \wedge$$

Finally, let  $L_{\text{loc}}$  denote the  $d$ -language over  $\Gamma$  defined by the first-order sentence  $\varphi_{\text{loc}}$  obtained by replacing each  $\Theta_\epsilon$  by the new unary relation symbol  $Q_\epsilon$  in the first-order part of  $\varphi'$ . In other words,  $\text{pixel}^d(L_{\text{loc}})$  is defined by the following first-order sentence:

$$\varphi_{\text{loc}} = \forall x \bigwedge_{i \in [d]} \left\{ \begin{array}{ll} \min_i(x) & \rightarrow \bigvee_{(\sharp, \epsilon) \in \Delta_i} Q_\epsilon(x) \\ \max_i(x) & \rightarrow \bigvee_{(\epsilon, \sharp) \in \Delta_i} Q_\epsilon(x) \\ \neg \max_i(x) & \rightarrow \bigvee_{(\epsilon, \epsilon') \in \Delta_i} (Q_\epsilon(x) \wedge Q_{\epsilon'}(\text{succ}_i(x))) \end{array} \right\} \wedge$$

Hence,  $L_{\text{loc}} = L(\Delta_1, \dots, \Delta_d)$ . That is,  $L_{\text{loc}}$  is indeed local and the corresponding sets of tiles are the  $\Delta_i$ 's of the previous formula. It is now easy to see that our initial  $d$ -language  $L$  is the projection of the local language



$L_{\text{loc}}$  by the projection  $\pi : \Gamma \rightarrow \Sigma$  defined as follows:  $\pi(\epsilon) = s$  iff  $\epsilon_i = 1$  for  $i \in [m]$  and  $U_i$  is  $Q_s$ . This completes the proof.  $\square$

**Proposition 2.9.**  $\text{ESO}(\text{arity } 1) \subseteq \text{ESO}(\forall^1, \text{arity } 1)$  on pixel structures, for any  $d > 0$ .

*Sketch of Proof.* In a pixel structure, each function symbol  $\text{succ}_i$  is interpreted as a cyclic successor, that is a bijective function. So, a pixel structure is a *bijective structure*, that is a first-order unary structure whose (unary) functions  $f$  are bijective and that explicitly includes all their inverse bijections  $f^{-1}$ . It has been proved in [DG06] that any first-order formula on a bijective structure can be rewritten as a so-called *cardinality formula*, that is as a boolean combination of sentences of the form  $\psi^{\geq k} = \exists^{\geq k} x \psi(x)$  (for  $k \geq 1$ ) where  $\psi(x)$  is a quantifier-free formula with the single variable  $x$  and where the quantifier  $\exists^{\geq k} x$  means "there exist at least  $k$  elements  $x$ ". Therefore, it is easily seen that proving the proposition amounts to show that each sentence of the form  $\psi^{\geq k}$  or  $\neg\psi^{\geq k}$  can be translated in  $\text{ESO}(\forall^1, \text{arity } 1)$  on pixel structures.

This is done as follows: for a given sentence  $\exists^{\geq k} x \psi(x)$ , we introduce new unary relations  $U^{=0}, U^{=1}, \dots, U^{=k-1}$  and  $U^{\geq k}$ , with the intended meaning:

A pixel  $a \in [n]^d$  belongs to  $U^{=j}$  (resp.  $U^{\geq k}$ ) iff there are exactly  $j$  (resp. at least  $k$ ) pixels  $b \in [n]^d$  lexicographically smaller than or equal to  $a$  such that  $\text{pixel}^d(p) \models \psi(b)$ .

Then we have to compel these relation symbols to fit their expected interpretations, by means of a first-order formula with a single universally quantified variable. First, we demand the relations to be pairwise disjoint:

$$(1) \bigwedge_{i < j < k} \left( \neg U^{=i}(x) \vee \neg U^{=j}(x) \right) \wedge \bigwedge_{i < k} \left( \neg U^{=i}(x) \vee \neg U^{\geq k}(x) \right).$$

Then, we temporarily denote by  $\leq_{\text{lex}}$  the lexicographic order on  $[n]^d$  inherited from the natural order on  $[n]$ , and by  $\text{succ}_{\text{lex}}$ ,  $\text{min}_{\text{lex}}$ ,  $\text{max}_{\text{lex}}$  its associated successor function and unary relations corresponding to extremal elements. Then the sets described above can be defined inductively by the conjunction of the following six formulas:

$$(2) (\text{min}_{\text{lex}}(x) \wedge \neg\psi(x)) \rightarrow U^{=0}(x)$$

$$(3) (\text{min}_{\text{lex}}(x) \wedge \psi(x)) \rightarrow U^{=1}(x)$$

$$(4) \bigwedge_{i < k} \left( \neg \text{max}_{\text{lex}}(x) \wedge U^{=i}(x) \wedge \neg\psi(\text{succ}_{\text{lex}}(x)) \right) \rightarrow U^{=i}(\text{succ}_{\text{lex}}(x))$$

$$(5) \bigwedge_{i < k-1} \left( \neg \text{max}_{\text{lex}}(x) \wedge U^{=i}(x) \wedge \psi(\text{succ}_{\text{lex}}(x)) \right) \rightarrow U^{=i+1}(\text{succ}_{\text{lex}}(x))$$

$$(6) \left( \neg \text{max}_{\text{lex}}(x) \wedge U^{=k-1}(x) \wedge \psi(\text{succ}_{\text{lex}}(x)) \right) \rightarrow U^{\geq k}(\text{succ}_{\text{lex}}(x))$$

$$(7) \left( \neg \text{max}_{\text{lex}}(x) \wedge U^{\geq k}(x) \right) \rightarrow U^{\geq k}(\text{succ}_{\text{lex}}(x)).$$

Hence, under the hypothesis  $(1) \wedge \dots \wedge (7)$ , the sentences  $\psi^{\geq k}$  and  $\neg\psi^{\geq k}$  are equivalent, respectively, to  $\forall x(\max_{\text{lex}}(x) \rightarrow U^{\geq k}(x))$  and  $\forall x(\max_{\text{lex}}(x) \rightarrow \neg U^{\geq k}(x))$ .

To complete the proof, it remains to get rid of the symbols  $\text{succ}_{\text{lex}}$ ,  $\min_{\text{lex}}$  and  $\max_{\text{lex}}$  that are not allowed in our language. It is done by referring to these symbols implicitly rather than explicitly. For instance, since  $\text{succ}_{\text{lex}}(x) = \text{succ}_i \text{succ}_{i+1} \dots \text{succ}_d(x)$ , for each non maximal  $x \in [n]^d$ , i.e., distinct from  $(n, \dots, n)$ , and for the smallest  $i \in [d]$  such that  $\bigwedge_{j>i} \max_j(x)$ , each formula  $\varphi$  involving  $\text{succ}_{\text{lex}}(x)$  actually corresponds to the conjunction:

$$\bigwedge_{i \in [d]} \left( \neg \max_i(x) \wedge \bigwedge_{i < j \leq d} \max_j(x) \rightarrow \varphi_i \right),$$

where  $\varphi_i$  is obtained from  $\varphi$  by the substitution  $\text{succ}_{\text{lex}}(x) \rightsquigarrow \text{succ}_i \text{succ}_{i+1} \dots \text{succ}_d(x)$ . Similar arguments allow to get rid of  $\min_{\text{lex}}$  and  $\max_{\text{lex}}$ .  $\square$

**Remark 2.10.** *In this proof, two crucial features of a structure of type  $\text{pixel}^d(p)$  are involved:*

- its "bijective" nature, that allows to rewrite first-order formulas as cardinality formulas with a single first-order variable;
- the "regularity" of its predefined arithmetics (the functions  $\text{succ}_i$  defined on each dimension), that endows  $\text{pixel}^d(p)$  with a grid structure: it enables us to implicitly define a linear order of the whole domain  $\text{dom}(p)$  by means of first-order formulas with a single variable, which in turn allows to express cardinality formulas by "cumulative" arguments, via the sets  $U^i$ .

Proposition 3.9 straightforwardly generalizes to all structures – and there are a lot – that fulfill these two properties.

To conclude this section, let us mention that we could rather easily derive from Theorem 2.6 the following additional characterization of  $\text{REC}^d$ :

**Corollary 2.11.** *For any  $d > 0$  and any  $d$ -language  $L$ , the following assertions are equivalent:*

1.  $L \in \text{REC}^d$ ;
2.  $\text{pixel}^d(L) \in \text{ESO}(\text{var } 1)$ .

### 3. A logical characterization of $\text{NLIN}_{\text{ca}}$

Besides the notion of recognizable picture language, the main concept studied in this paper is the linear time complexity on nondeterministic cellular automaton of any dimension.

**Definition 3.1.** A one-way  $d$ -dimensional cellular automaton ( $d$ -automaton, for short) over an alphabet  $\Sigma$  is a tuple  $\mathcal{A} = (\Sigma, \Gamma, \delta, F)$ , where

- the finite alphabet  $\Gamma$  called the **set of states** of  $\mathcal{A}$  includes the **input alphabet**  $\Sigma$  and the set  $F$  of **accepting states**:  $\Sigma, F \subseteq \Gamma$ ;
- $\delta$  is the (nondeterministic) **transition function** of  $\mathcal{A}$ :  $\delta : \Gamma \times (\Gamma^\#)^d \rightarrow \mathcal{P}(\Gamma)$ .

**Definition 3.2.** Let  $\mathcal{A} = (\Sigma, \Gamma, \delta, F)$  be a  $d$ -automaton and  $p, p' : [n]^d \rightarrow \Gamma$  be two  $d$ -pictures on  $\Gamma$ . We say that  $p'$  is a **successor** of  $p$  for  $\mathcal{A}$ , denoted by  $p' \in \mathcal{A}(p)$ , if for each  $a \in [n]^d$ ,

$$p'(a) \in \delta(p(a), p^\sharp(a^{(1)}), \dots, p^\sharp(a^{(d)})),$$

where  $p^\sharp$  denotes the bordered picture of  $p$  and, for  $i \in [d]$ ,  $a^{(i)}$  denotes the point having the same coordinates as point  $a$  except the  $i^{\text{th}}$  one which is  $a_i + 1$ . So, the value of  $p'$  at point  $a \in [n]^d$  only depends on the set of points  $a, a^{(1)}, \dots, a^{(d)}$  of  $p$ . This set of  $d+1$  points is called the **neighborhood** of  $a$ .

The set of  $j^{\text{th}}$ -successors of  $p$  for  $\mathcal{A}$ , denoted  $\mathcal{A}^j(p)$ , is defined inductively:

$$\mathcal{A}^0(p) = \{p\} \text{ and for } j \geq 0: \mathcal{A}^{j+1}(p) = \bigcup_{p' \in \mathcal{A}^j(p)} \mathcal{A}(p').$$

**Definition 3.3.** A **computation** of a  $d$ -automaton  $\mathcal{A} = (\Sigma, \Gamma, \delta, F)$  on an input  $d$ -picture  $p$  on  $\Sigma$  is a sequence  $p_1, p_2, p_3, \dots$  of  $d$ -pictures on  $\Gamma$  such that  $p_1 = p$  and  $p_{i+1} \in \mathcal{A}(p_i)$  for each  $i$ . The picture  $p_i$ ,  $i \geq 1$ , is called the  $i^{\text{th}}$  **configuration** of the computation. A computation is **accepting** if it is finite – it has the form  $p_1, p_2, \dots, p_k$  for some  $k$  – and the cell of minimal coordinates,  $1^d = (1, \dots, 1)$ , of its last configuration is in an accepting state:  $p_k(1^d) \in F$ .

**Remark 3.4.** Note that the space used by a  $d$ -automaton is exactly the space (set of cells) occupied by its input  $d$ -picture.

**Definition 3.5.** Let  $\mathcal{A} = (\Sigma, \Gamma, \delta, F)$  be a  $d$ -automaton and let  $T : \mathbb{N} \rightarrow \mathbb{N}$  be a function such that  $T(n) > n$ . A  $d$ -picture  $p : [n]^d \rightarrow \Sigma$  is **accepted by  $\mathcal{A}$  in time  $T$**  if  $\mathcal{A}$  admits an accepting computation of length  $T(n)$  on  $p$ . That means, there exists a computation  $p = p_1, p_2, \dots, p_{T(n)}$  of  $\mathcal{A}$  on  $p$  such that  $p_{T(n)}(1^d) \in F$ .

A  $d$ -language  $L$  on  $\Sigma$  is **accepted**, or **recognized**, by  $\mathcal{A}$  in time  $T$  if it is the set of  $d$ -pictures accepted by  $\mathcal{A}$  in time  $T$ . That is

$$L = \{p : \exists p' \in \mathcal{A}^{T(n)-1}(p) \text{ such that } p'(1^d) \in F \text{ where } \text{dom}(p) = [n]^d\}.$$

If  $T(n) = cn + c'$ , for some integers  $c, c'$ , then  $L$  is said to be **recognized in linear time** and we write  $L \in \text{NLIN}_{\text{ca}}^d$ .

The time bound  $T(n) > n$  of the above definition is necessary and sufficient to allow the information of any pixel of  $p$  to be communicated to the pixel of minimal coordinates,  $1^d$ .

**Remark 3.6.** The nondeterministic linear time class  $\text{NLIN}_{\text{ca}}^d$  is very robust, i.e. is not modified by many changes in the definition of the automaton or in its time bound. In particular, the constants  $c, c'$  defining the bound  $T(n) = cn + c'$  can be fixed arbitrarily, provided  $T(n) > n$ . Typically, the class  $\text{NLIN}_{\text{ca}}^d$  does not change if we take the minimal time  $T(n) = n + 1$ , called real time.

As we have stated several logical characterizations of the class  $\text{REC}^d$  of recognizable picture languages (Theorem 2.6), we now give two characterizations of the class  $\text{NLIN}_{\text{ca}}^d$  in existential second-order logic.

**Theorem 3.7.** For any  $d > 0$  and any  $d$ -language  $L$ , the following assertions are equivalent:

1.  $L \in \text{NLIN}_{\text{ca}}^d$ ;
2.  $\text{coord}^d(L) \in \text{ESO}(\forall^{d+1}, \text{arity } d+1)$ ;

3.  $\text{coord}^d(L) \in \text{ESO}(\text{var } d + 1)$ .

This theorem is a straightforward consequence of Proposition 3.8 and 3.9 below. The former states the equivalence (1)  $\Leftrightarrow$  (2); with the latter, we establish the normalization  $\text{ESO}(\text{var } d) \subseteq \text{ESO}(\forall^{d+1}, \text{arity } d)$  on coordinate structures, which yields equivalence (2)  $\Leftrightarrow$  (3).

**Proposition 3.8.** *For any  $d > 0$  and any  $d$ -language  $L$ ,*

$$L \in \text{NLIN}_{ca}^d \Leftrightarrow \text{coord}^d(L) \in \text{ESO}(\forall^{d+1}, \text{arity } d + 1).$$

*Sketch of Proof.*  $\Rightarrow$  Let  $L \in \text{NLIN}_{ca}^d$ . By Remark 3.6, let us assume without loss of generality that  $L$  is recognized by a  $d$ -automaton  $\mathcal{A} = (\Sigma, \Gamma, \delta, F)$  in time  $n + 1$ . The sentence in  $\text{ESO}(\forall^{d+1}, \text{arity } d + 1)$  that we construct is of the form  $\exists(R_s)_{s \in \Gamma} \forall \mathbf{x} \psi(\mathbf{x})$ , where  $\psi(\mathbf{x})$  is a quantifier-free formula such that:

- $\psi$  uses a list of exactly  $d + 1$  first-order variables  $\mathbf{x} = (x_1, \dots, x_d, x_{d+1})$ . Intuitively, the  $d$  first ones represent the coordinates of any point in  $\text{dom}(p) = [n]^d$  and the variable  $x_{d+1}$  represents any of the first  $n$  instants  $t \in [n]$  of the computation (the last instant  $n + 1$  is not explicitly represented);
- $\psi$  uses, for each state  $s \in \Gamma$ , a relation symbol  $R_s$  of arity  $d + 1$ . Intuitively,  $R_s(a_1, \dots, a_d, t)$  holds, for any  $a = (a_1, \dots, a_d) \in [n]^d$  and any  $t \in [n]$ , iff the state of cell  $a$  at instant  $t$  is  $s$ .

$\psi(\mathbf{x})$  is the conjunction  $\psi(\mathbf{x}) = \text{INIT}(\mathbf{x}) \wedge \text{STEP}(\mathbf{x}) \wedge \text{END}(\mathbf{x})$  of three formulas whose intuitive meaning is the following.

- $\forall \mathbf{x} \text{INIT}(\mathbf{x})$  describes the first configuration of  $\mathcal{A}$ , i.e. at initial instant 1;
- $\forall \mathbf{x} \text{STEP}(\mathbf{x})$  describes the computation of  $\mathcal{A}$  between the instants  $t$  and  $t + 1$ , for  $t \in [n - 1]$ , i.e. describes the  $(t + 1)^{\text{th}}$  configuration from the  $t^{\text{th}}$  one;
- $\forall \mathbf{x} \text{END}(\mathbf{x})$  expresses that the  $n^{\text{th}}$  configuration of  $\mathcal{A}$  leads to a (last)  $(n + 1)^{\text{th}}$  configuration which is accepting, i.e. with an accepting state in cell  $1^d$ .

Let us give explicitly the main formula  $\text{STEP}$ . The main technical difficulty comes from the description of the transition function  $\delta$  of  $\mathcal{A}$  for any cell in the border of the picture. In order to describe uniformly  $\delta$  we introduce the notion of *neighborhood*. A *neighborhood* for  $\mathcal{A}$  is any function  $\nu : [0, d] \rightarrow \Gamma^\#$  such that  $\nu(0) \in \Gamma$ . Let  $\text{neighb}(\mathcal{A})$  denote the set of neighborhoods  $\nu = (\nu(0), \nu(1), \dots, \nu(d))$  of  $\mathcal{A}$ . So,  $\delta$  is a function of prototype  $\text{neighb}(\mathcal{A}) \rightarrow \mathcal{P}(\Gamma)$ .

Clearly, the universal closure  $\forall \mathbf{x} \text{STEP}(\mathbf{x})$  of the following formula  $\text{STEP}(\mathbf{x})$  — that uses the "neighborhood" notation — correctly describes the computation between the instants  $t$  and  $t + 1$ , for  $t \in [n - 1]$ :

$$\text{STEP}(\mathbf{x}) = \bigwedge_{\nu \in \text{neighb}(\mathcal{A})} \left\{ \begin{array}{l} \neg \max(x_{d+1}) \wedge R_{\nu(0)}(\mathbf{x}) \\ \bigwedge_{i \in [d]: \nu(i) \in \Gamma} \left( \neg \max(x_i) \wedge R_{\nu(i)}(\mathbf{x}^{(i)}) \right) \\ \bigwedge_{i \in [d]: \nu(i) = \#} \max(x_i) \end{array} \right\} \wedge \bigoplus_{s \in \delta(\nu)} R_s(\mathbf{x}^{(d+1)}).$$

Here,  $\bigoplus$  denotes the exclusive disjunction and  $\mathbf{x}^{(i)}$  is the  $(d + 1)$ -tuple  $\mathbf{x}$  where  $x_i$  is replaced by  $\text{succ}(x_i)$ . Hence we have proved that for any  $d$ -picture  $p$  on  $\Sigma$ ,  $\text{coord}^d(p)$  satisfies the  $\text{ESO}(\forall^{d+1}, \text{arity } d + 1)$ -sentence  $\exists(R_s)_{s \in \Gamma} \forall \mathbf{x} \psi(\mathbf{x})$  if, and only if,  $\mathcal{A}^n(p)$  contains an accepting configuration, i.e.  $\mathcal{A}$  accepts  $p$  in time  $n + 1$ , or, by definition,  $p \in L$ . Hence  $L \in \text{ESO}(\forall^{d+1}, \text{arity } d + 1)$ , as required.

◁ Assume  $\text{coord}^d(L) \in \text{ESO}(\forall^{d+1}, \text{arity } d+1)$ , i.e. there is some sentence  $\Phi \in \text{ESO}(\forall^{d+1}, \text{arity } d+1)$  such that  $p \in L \Leftrightarrow \text{coord}^d(p) \models \Phi$ . We want to prove  $L \in \text{NLIN}_{\text{ca}}^d$ , i.e.  $L$  is recognized by some  $d$ -automaton in linear time. Let us give the main idea of the proof for the simplest case  $d = 1$  and a formula  $\Phi \in \text{ESO}(\forall^2, \text{arity } 2)$  of the form

$$\Phi = \exists R \forall x \forall y \psi(x, y)$$

where  $R$  is a binary relation symbol and  $\psi$  is a quantifier-free formula where the only atoms in which  $R$  occurs, called  $R$  atoms, are of the following forms (1-4):

$$(1) R(x, y); \quad (2) R(\text{succ}(x), y); \quad (3) R(x, \text{succ}(y)); \quad (4) R(y, x).$$

First, notice that if the only atoms where  $R$  occurs are of the forms (1-3), i.e. the variables  $x, y$  only appear in this *unique* order in the arguments of  $R$ , then formula  $\Phi$  has a *local behaviour*: points  $(x, y)$ ,  $(\text{succ}(x), y)$  and  $(x, \text{succ}(y))$  are *neighbours*, i.e. adjacent each other. This allows to construct a 1-automaton (nondeterministic cellular automaton of dimension 1)  $\mathcal{A}$  that "mimics"  $\Phi$ . Roughly,  $\mathcal{A}$  "guesses" successively "rows"  $R(i, \dots)$ , for  $i = 1, 2, \dots, n$ , of  $R$ , and in the same time, it "checks locally" the coherence of each "instantiation"  $\psi(i, j)$ : more precisely, at "instant"  $i$ , the state of each cell  $j$ ,  $1 \leq j \leq n$ , of  $\mathcal{A}$  "contains" both values  $R(i, j)$  and  $R(i+1, j)$ . So, in case  $R$ -atoms are of the forms (1-3), the language  $L$  is recognized by such a 1-automaton  $\mathcal{A}$  in linear time as claimed.

Now, let us consider the "general" case where the formula includes all the forms (1-4). Of course, the "pixel" (4)  $R(y, x)$  is not adjacent to "pixels" (1-3) but it is their symmetric (more precisely, it is the symmetric of  $R(x, y)$ ) with respect to the "diagonal"  $x = y$ . The intuitive idea is "to cut" or "to fold" the "picture"  $R$  along this diagonal:  $R$  is replaced by its two "half pictures" denoted  $R_1$  and  $R_2$ , that are "superposed" in the "half square"  $x \leq y$  above the diagonal. More precisely,  $R_1$  and  $R_2$  are binary relations whose intuitive meaning is the following: for points  $(x, y)$  such that  $x \leq y$ , one has the equivalence  $R_1(x, y) \Leftrightarrow R(x, y)$  and the equivalence  $R_2(x, y) \Leftrightarrow R(y, x)$ . By this transformation, each pixel  $R_2(x, y)$  that represents the original pixel  $R(y, x)$  lies at the same point  $(x, y)$  as pixel  $R_1(x, y)$  that represents pixel  $R(x, y)$ , for  $x \leq y$  (the case  $y \leq x$  is similar). This solves the problem of vicinity.

More precisely, the sentence  $\Phi = \exists R \forall x \forall y \psi(x, y)$  is normalized as follows. Let  $\text{coherent}(x, y)$  denote the formula

$$x = y \rightarrow (R_1(x, y) \leftrightarrow R_2(x, y))$$

whose universal closure ensures the coherence of  $R_1$  and  $R_2$  on the common part of  $R$  they both represent, that is the diagonal  $x = y$ . Using  $R_1$  and  $R_2$ , it is not difficult to construct a formula

$$\psi'(x, y) = \left( \begin{array}{l} \text{coherent}(x, y) \quad \wedge \\ x < y \rightarrow \psi_{<}(x, y) \quad \wedge \\ x = y \rightarrow \psi_{=}(x, y) \quad \wedge \\ x > y \rightarrow \psi_{>}(x, y) \end{array} \right)$$

such that the sentence  $\Phi' = \exists R_1 \exists R_2 \forall x \forall y \psi'(x, y)$  in  $\text{ESO}(\forall^2, \text{arity } 2)$  is equivalent to  $\Phi$ . Let us describe and justify its precise form and meaning.

The formulas  $\psi_{<}(x, y)$ ,  $\psi_{=}(x, y)$  and  $\psi_{>}(x, y)$  are obtained from formula  $\psi(x, y)$  by substitution of  $R$  atoms by  $R_1$  or  $R_2$  atoms according to the cases described in Table 1. It is easy to check that each replacement is correct according to its case. For instance, it is justified to replace each atom of the form  $R(x, \text{succ}(y))$  in  $\psi$  by  $R_2(\text{succ}(y), x)$  when  $x > y$  (in order to obtain the formula  $\psi_{>}(x, y)$ ) because when  $x > y$ , then  $\text{succ}(y) \leq x$  and hence the equivalence  $R(x, \text{succ}(y)) \leftrightarrow R_2(\text{succ}(y), x)$  holds, by definition of  $R_2$ .

Notice that the variables  $x, y$  always occur in this order in each  $R_1$  or  $R_2$  atom of the formulas  $\psi_{<}$  and  $\psi_{=}$  (see Table 1). At the opposite, they always occur in the reverse order  $y, x$  in the formula  $\psi_{>}(x, y)$ . This

Table 1: Replacement of  $R$ -atoms by  $R_1$ - or  $R_2$ -atoms

case	formula	$R(x, y)$	$R(\text{succ}(x), y)$	$R(x, \text{succ}(y))$	$R(y, x)$
$x < y$	$\psi_{<}(x, y)$	$R_1(x, y)$	$R_1(\text{succ}(x), y)$	$R_1(x, \text{succ}(y))$	$R_2(x, y)$
$x = y$	$\psi_{=}(x, y)$	$R_1(x, y)$	$R_2(x, \text{succ}(y))$	$R_1(x, \text{succ}(y))$	$R_1(x, y)$
$x > y$	$\psi_{>}(x, y)$	$R_2(y, x)$	$R_2(y, \text{succ}(x))$	$R_2(\text{succ}(y), x)$	$R_1(y, x)$

is not a problem because, by symmetry, the roles of  $x$  and  $y$  can be exchanged and the universal closure  $\forall x \forall y (x > y \rightarrow \psi_{>}(x, y))$  is trivially equivalent to  $\forall x \forall y (y > x \rightarrow \psi_{>}(y, x))$ . So, the above sentence  $\Phi'$  – and hence, the original sentence  $\Phi$  – is equivalent to the sentence denoted  $\Phi''$  obtained by replacing in  $\Phi'$  the subformula  $x > y \rightarrow \psi_{>}(x, y)$  by  $y > x \rightarrow \psi_{>}(y, x)$ . By construction, relation symbols  $R_1, R_2$  only occur in  $\Phi$  in atoms of the three required "sorted" forms:  $R_i(x, y)$ ,  $R_i(\text{succ}(x), y)$  or  $R_i(x, \text{succ}(y))$ . Finally, to be precise, there remain two difficulties so that a 1-automaton can simulate the "sorted" sentence  $\Phi''$  in linear time, by the informal algorithm described above:

- the presence of equalities and inequalities in the sentence;
- the forms of the atoms involving input relation symbols.

It is easy to get rid of equalities and inequalities by introducing new binary relation symbols defined and used in a "sorted" manner too (see (1-3)). Concerning the second point, we can assume, without loss of generality, that the only atoms involving the input relation symbols  $(Q_s)_{s \in \Sigma}$  are of the two forms  $Q_s(x)$  or  $Q_s(y)$ . As we do for equalities and inequalities, we can get rid of atoms of the form  $Q_s(y)$  by introducing new binary ESO relation symbols : intuitively, they convey each bit  $Q_s(a)$  at each point of coordinates  $(a, \dots)$  or  $(\dots, a)$ ; those new binary relations are also defined and used in a "sorted" manner. The fact that all the atoms involving the input are of the form  $Q_s(x)$  allows to consider this input in the initial configuration of the computation of the 1-automaton *but in no later configuration* as required. So, the sketch of proof is complete for the case  $d = 1$ .

For the general case, i.e. for any dimension  $d$ , the ideas and the steps of the proof are exactly the same as for  $d = 1$  but the notations and details of the proof are much more technical. To give an idea, let us succinctly describe the ESO relations of arity  $d + 1$  introduced in the main normalization step. Here again, each ESO relation symbol  $R$  of the original sentence  $\Phi$  in  $\text{ESO}(\forall^{d+1}, \text{arity } d + 1)$  is replaced by – or, intuitively, "divided into" –  $(d + 1)!$  new ESO relation symbol  $R_\alpha$  of the same arity  $d + 1$ , where  $\alpha$  is a permutation of the set of indices  $[d + 1]$ . The intended meaning of each relation  $R_\alpha$  is the following: for each tuple  $(a_1, \dots, a_{d+1}) \in [n]^{d+1}$  such that  $a_1 \leq a_2 \leq \dots \leq a_{d+1}$ , the equivalence

$$R_\alpha(a_1, \dots, a_{d+1}) \leftrightarrow R(a_{\alpha(1)}, \dots, a_{\alpha(d+1)})$$

holds. Then, we introduce a partition of the domain  $[n]^{d+1}$  into subdomains, similar to the partition of the domain  $[n]^2$  described above for  $d = 1$  into the diagonal  $x = y$  and the two "half domains" over and under the diagonal  $x < y$  and  $x > y$ , respectively. According to the case (i.e. subdomain of the partition), this allows to replace each  $R$  atom in  $\Phi$  by an atom of one of the two following "sorted" forms:

$$R_\alpha(\mathbf{x}); R_\alpha(\mathbf{x}^{(i)})$$

where  $\mathbf{x} = (x_1, \dots, x_{d+1})$ ,  $1 \leq i \leq d + 1$ , and  $\mathbf{x}^{(i)}$  is the tuple  $\mathbf{x}$  where  $x_i$  is replaced by  $\text{succ}(x_i)$ . Finally, the equalities and inequalities are similarly eliminated in the sentence and we normalize it with respect

to the input  $d$ -ary relations  $(Q_s)_{s \in \Sigma}$  by using new ESO relation symbols of arity  $d + 1$  to convey the input information: in the final "sorted" sentence all the  $Q_s$  atoms are of the unique form  $Q_s(x_1, \dots, x_d)$ . For such a "sorted"  $\text{ESO}(\forall^{d+1}, \text{arity } d + 1)$ -sentence  $\Phi$ , it is now easy to construct a  $d$ -automaton that generalizes the automaton described above in case  $d = 1$ , and checks in linear time whether  $\text{coord}^d(p) \models \Phi$ .  $\square$

**Proposition 3.9.** *For any  $d > 0$ ,  $\text{ESO}(\text{var } d) \subseteq \text{ESO}(\forall^d, \text{arity } d)$  on coordinate structures.*

*Sketch of Proof.* We first prove a kind of Skolemization of  $\text{ESO}(\text{var } d)$ -formulas, thus providing a first normalization of these formulas, in which the first-order part is *universal* and includes *the same number of first-order variables* than the initial formula. To illustrate the procedure that performs this preliminary normalization, let us run it on a very simple first-order formula with *two* variables:

$$\varphi \equiv \exists x (\forall y U(x, y) \vee \exists y D(x, y)).$$

We introduce three new relation symbols  $R_1, R_2, R_3$  corresponding to the quantified subformulas of  $\varphi$ .

$$\begin{aligned} \text{def}_1(R_1) &\equiv \forall x : R_1(x) \leftrightarrow \forall y U(x, y) \\ \text{def}_2(R_2) &\equiv \forall x : R_2(x) \leftrightarrow \exists y D(x, y) \\ \text{def}_3(R_3) &\equiv \forall x : R_3(x) \leftrightarrow R_1(x) \vee R_2(x) \end{aligned}$$

Hence, our initial formula can be rewritten:

$$\exists R_1, R_2, R_3 : \left( \bigwedge_{1 \leq i \leq 3} \text{def}_i(R_i) \right) \wedge \exists x R_3(x). \quad (4)$$

It is easily seen that (4) can be written as a conjunction of prenex formulas, each of which involves no more than two variables and has a quantifier prefix of the shape  $\forall x \forall y$  or  $\forall x \exists y$  (we include in this latter form the subformula  $\exists x R_3(x)$ ). All in all,  $\varphi$  is equivalent to a formula of the form:

$$\exists R_1, R_2, R_3 : \forall x \forall y \psi(x, y, \mathbf{R}) \wedge \forall x \exists y \theta(x, y, \mathbf{R}), \quad (5)$$

where  $\psi$  and  $\theta$  are quantifier-free. In order to put this conjunction under prenex form without adding a new first-order variable, we have to "replace" the existential quantifier by a universal one. (Afterward  $\varphi$ , as a conjunction of formulas of prefix  $\forall x, y$ , could be written under the requisite shape.) To proceed, we get use of the arithmetics embedded in coordinate structures. It allows to defining a binary relation  $W$  with intended meaning:  $W(x, y)$  iff there exists  $z \leq y$  such that  $\theta(x, z)$  holds. This interpretation is achieved thanks to the formula:

$$\forall x, y \left\{ \begin{array}{l} \min(y) \rightarrow (W(x, y) \leftrightarrow \theta(x, y)) \quad \wedge \\ \neg \max(y) \rightarrow (W(x, \text{succ}(y)) \leftrightarrow (\theta(x, \text{succ}(y)) \vee W(x, y))) \end{array} \right\} \quad (6)$$

Under assumption (6), the assertion  $\forall x \exists y \theta(x, y)$  is equivalent to  $\forall x \forall y : \max(y) \rightarrow W(x, y)$ . This allows to rewrite (5), and hence  $\varphi$ , as the formula:

$$\exists R_1, R_2, R_3, W \left( \begin{array}{l} (6) \wedge \forall x \forall y \psi(x, y, \mathbf{R}) \wedge \\ \forall x \forall y (\max(y) \rightarrow W(x, y)) \end{array} \right) \quad (7)$$

which belongs to  $\text{ESO}(\forall^2)$ .

Thus, the above considerations allow to show the normalization  $\text{ESO}(\text{var } d) = \text{ESO}(\forall^d)$  on coordinate structures. It remains to prove  $\text{ESO}(\forall^d) = \text{ESO}(\forall^d, \text{arity } d)$ . It amounts to build, for each formula  $\Phi$  of type

$\exists \mathbf{R} \forall x_1, \dots, x_d \varphi$ , where  $\varphi$  is quantifier-free and  $\mathbf{R}$  is a tuple of relation symbols of any arity, a formula  $\Phi'$  with the same shape, but in which all relation symbols are of arity  $\leq d$ , such that  $\Phi$  and  $\Phi'$  have the same models, as far as coordinate structures are concerned. The possibility to replace a  $k$ -ary ( $k \geq d$ ) relation symbol  $R$  of  $\Phi$  by  $d$ -ary symbols rests in the limitation of the number of first-order variables in  $\Phi$ : each atomic formula involving  $R$  has the form  $R(t_1, \dots, t_k)$  where the  $t_i$ 's are terms built on  $x_1, \dots, x_d$ . Therefore, although  $R$  is  $k$ -ary, *in each of its occurrences* it behaves as a  $d$ -ary symbol, dealing with the sole variables  $x_1, \dots, x_d$ . Hence, the key is to create a  $d$ -ary symbol for each occurrence of  $R$  in  $\Phi$  or, more precisely, for each  $k$ -tuple of terms  $(t_1, \dots, t_k)$  involved in a  $R$ -atomic formula. Let us again opt for a “proof-by-example” choice and illustrate the procedure on a very simple case.

Let  $\Phi$  be the  $\text{ESO}(\forall^2, \text{arity } 3)$ -formula  $\exists R \forall x, y \varphi(x, y, R)$ , where  $\varphi \equiv R(x, y, x) \wedge \neg R(y, x, y)$ . Introduce two new *binary* relation symbols  $R_{(x,y,x)}$  and  $R_{(y,x,y)}$  associated to the triple of terms  $(x, y, x)$  and  $(y, x, y)$  involved in  $\Phi$ , and fix their interpretation as follows: for any  $a, b \in [n]$ ,

$$R_{(x,y,x)}(a, b) \Leftrightarrow R(a, b, a) \text{ and } R_{(y,x,y)}(a, b) \Leftrightarrow R(b, a, b).$$

Then we get the equivalence:

$$\langle S, R \rangle \models \forall x, y : R(x, y, x) \wedge \neg R(y, x, y) \Leftrightarrow \langle S, \mathbf{R} \rangle \models \forall x, y : R_{(x,y,x)}(x, y) \wedge \neg R_{(y,x,y)}(x, y)$$

which, in turn, yields the implication:

$$S \models \exists R \forall x, y : R(x, y, x) \wedge \neg R(y, x, y) \Rightarrow S \models \exists \mathbf{R} \forall x, y : R_{(x,y,x)}(x, y) \wedge \neg R_{(y,x,y)}(x, y) \quad (8)$$

The converse implication would immediately complete the proof. Unfortunately, it does not hold, since the second formula has a model, while the first has not.

To get the converse (right-to-left) implication of (8), we have to strengthen the second formula with some assertion that compels the tuple  $R_{(x,y,x)}, R_{(y,x,y)}$  to be, in some sense, the binary representation of some ternary relation. This last construction is more sophisticated than the preceding ones, and we can't detail it here.  $\square$

#### 4. Hierarchy results

In Section 2, we have established several natural characterizations of the classes  $\text{REC}^d$  in existential second-order logic on pixel structures (Theorem 2.6 and Corollary 2.11). Is there similar characterizations in coordinate structures instead of pixel structures? One can trivially translate the logical characterization of  $\text{REC}^d$  in  $\text{ESO}(\forall^1, \text{arity } 1)$  for the pixel representation into a characterization in  $\text{ESO}(\forall^d, \text{arity } d)$  for the coordinate representation. However, one notices that the sentences so obtained in  $\text{ESO}(\forall^d, \text{arity } d)$  are “sorted” in a sense similar to that evoked in the previous section (see the proof of Proposition 3.8): the atoms involving the ESO (resp. input) relation symbols, all of arity  $d$ , are of the “sorted” form  $R(\mathbf{x})$  or  $R(\mathbf{x}^{(i)})$  (resp.  $Q_s(\mathbf{x})$ ) where  $\mathbf{x}$  is the “sorted” tuple of variables  $(x_1, \dots, x_d)$  and  $\mathbf{x}^{(i)}$ ,  $i \in [d]$ , is the same “sorted” tuple where  $x_i$  is replaced by  $\text{succ}(x_i)$ . In fact, one can prove that a  $d$ -language  $L$  is recognizable iff  $\text{coord}^d(L)$  is defined by a “sorted” sentence in  $\text{ESO}(\forall^d, \text{arity } d)$ . But requiring the sentence to be sorted seems rather artificial. Can it be dropped? The following result shows it cannot.

**Proposition 4.1.** *For each integer  $d \geq 2$  and for  $d$ -languages represented by their coordinate structures, the following strict inclusions hold:*

$$\text{ESO}(\forall^{d-1}, \text{arity } d-1) = \text{ESO}(\text{var } d-1) \subsetneq \text{REC}^d \subsetneq \text{ESO}(\text{var } d) = \text{ESO}(\forall^d, \text{arity } d).$$



That means the following implications hold,

$$\text{coord}^d(L) \in \text{ESO}(\text{var } d - 1) \Rightarrow L \in \text{REC}^d \Rightarrow \text{coord}^d(L) \in \text{ESO}(\text{var } d),$$

but neither of their converses does.

The proof of Proposition 4.6 that we skip here uses a simple lemma about the so-called  $\text{Mirror}_d$  language.

**Definition 4.2.** For  $d \geq 2$ , we denote by  $\text{Mirror}_d$  the set of pictures  $p : [n]^d \rightarrow \{0, 1\}$  that fulfill  $p(a) = p(a_{12})$  for all  $a \in [n]^d$ . Here,  $a_{12}$  is the tuple  $a$  in which the first and second components have been exchanged (e.g.,  $(5, 3, 4, 6)_{12} = (3, 5, 4, 6)$ ).

**Lemma 4.3.** For each  $d \geq 2$ , we trivially have  $\text{coord}^d(\text{Mirror}_d) \in \text{ESO}(\text{var } d)$ . However,  $\text{Mirror}_d \notin \text{REC}^d$ .

The proof of Lemma 4.3 essentially amounts to a counting argument. It is a simple variant of the proof of a similar result in Giammarresi et al. [GRST96] and we skip it.

Finally, the hierarchy theorem below (Theorem 4.6) involves the number of universal first-order quantifiers. It uses the following "symmetric" language as a counterexample.

**Definition 4.4.** Let  $\text{SYM}_d$  be the set of pictures  $p : [n]^d \rightarrow \{0, 1\}$  that fulfill, for every  $a = (a_1, \dots, a_d) \in [n]^d$ :

1.  $p(a) = p(a_\alpha)$  for each permutation  $\alpha$  of  $[d]$ , where  $a_\alpha = (a_{\alpha(1)}, \dots, a_{\alpha(d)})$ ;
2.  $p(a) = p(a_{\text{sym}(i)})$  for all  $i \in [d]$ , where  $a_{\text{sym}(i)}$  denotes the tuple  $a$  whose  $i^{\text{th}}$  component  $a_i$  is replaced by its "symmetric value"  $n + 1 - a_i$ <sup>4</sup>.

In other words, the values  $p(a)$  are defined up to all possible permutations of coordinates and up to all possible symmetries  $a \mapsto a_{\text{sym}(i)}$ .

**Lemma 4.5.** For all  $d \geq 2$ , we have  $\text{coord}^d(\text{SYM}_d) \in \text{ESO}(\forall^{d+1}, \text{arity } d) \setminus \text{ESO}(\forall^d, \text{arity } d)$ .

Here is our last hierarchy theorem.

**Theorem 4.6.** For each integer  $d \geq 2$  and for  $d$ -languages represented by coordinate structures, the following (strict) inclusions hold:

$$\begin{aligned} \text{REC}^d &\subsetneq \text{ESO}(\text{var } d) = \text{ESO}(\forall^d, \text{arity } d) \\ &\quad \cap \\ &\quad \text{ESO}(\forall^{d+1}, \text{arity } d) \\ &\quad \cap \\ \text{NLIN}_{ca}^d &= \text{ESO}(\text{var } d + 1) = \text{ESO}(\forall^{d+1}, \text{arity } d + 1) \end{aligned}$$

The above strict inclusion  $\text{ESO}(\text{var } d) \subsetneq \text{ESO}(\forall^{d+1}, \text{arity } d)$  trivially yields the following result.

**Corollary 4.7.** For each integer  $d \geq 2$  and for  $d$ -languages represented by coordinate structures, we have the strict inclusion:  $\text{ESO}(\text{var } d) \subsetneq \text{ESO}(\text{arity } d)$ .

**Remark 4.8.** Corollary 4.7 surprisingly contrasts with the equality  $\text{ESO}(\text{var } 1) = \text{ESO}(\text{arity } 1)$  for  $d$ -languages represented by pixels structures, see Theorem 2.6 and Corollary 2.11.

---

<sup>4</sup>Because of Item 1, Item 2 is clearly equivalent to the same assertion for only  $i = 1$ .

### Other hierarchies

It is also natural to address the question whether the following "hierarchies" about second-order logic (SO) over picture languages are strict or not:

1. Is there a strict hierarchy of SO or ESO according to the *second-order quantifier alternation*?
2. Is there a strict hierarchy of the classes  $\text{ESO}(\text{arity } d)$  and  $\text{ESO}(\forall^k, \text{arity } d)$  according to the *number of ESO relation symbols*?

It is well-known that the answer to question 1 is negative for MSO on word languages and tree languages: on these classes of languages,  $\text{MSO} = \text{EMSO}$  holds. That is, the hierarchy collapses at its first level.

At the opposite, Matz and Thomas [MT97] have answered question 1 positively for MSO over 2-dimensional picture languages: the quantifier alternation is strict for MSO on 2-picture languages (in pixel representation). We strongly conjecture that their result still holds with a similar proof for any dimension  $d > 2$  and the pixel representation. We also conjecture that similar results hold in the coordinate representation: typically, we think that for  $d$ -languages with  $d \geq 2$ , the quantifier alternation of  $\text{SO}(\text{arity } d)$  gives a strict hierarchy. Unfortunately, we have no idea to prove it, even in the "minimal case", that is  $\text{ESO}(\text{arity } 2) \subsetneq \text{SO}(\text{arity } 2)$ , for 2-languages.

In contrast, the answer to question 2 is totally known and uniform for picture languages of any dimension. In all cases, the hierarchy collapses at its first level. More precisely, Thomas [Tho82] has established that every EMSO sentence over words is equivalent to a sentence whose monadic quantifier prefix consists of a single existential quantifier. Matz [Mat98] has proved the same result over 2-pictures in the pixel representation. The proof of Matz can be extended (with slight adaptations) to any dimension  $d$ , for both pixel and coordinate representations. In other words, in both representations, all the logical classes – essentially  $\text{ESO}(\text{arity } d)$  and  $\text{ESO}(\forall^k, \text{arity } d)$  – we have studied over picture languages of any dimension are not modified by the requirement there should be only *one* ESO relation symbol.

### Some remarks about the dimensions of pictures

In this paper, for sake of simplicity and uniformity, we have chosen to restrict the presentation of our results to "square" pictures, i.e. pictures of prototype  $p : [n]^d \rightarrow \Sigma$ . This may appear as a too strong requirement. In this section, we explain how our results about logical classes and their relationships with complexity classes  $\text{REC}^d$  and  $\text{NLIN}_{ca}^d$  can be extended to the "most general" picture languages, i.e. to sets of  $d$ -pictures of prototype  $p : [n_1] \times \dots \times [n_d] \rightarrow \Sigma$ . "Most general" means as much general as they make sense in the logical or complexity theoretical framework involved.

#### $\text{REC}^d$ and logical characterizations in pixel encodings

All our results about logical characterizations of picture languages in pixel encodings and the class  $\text{REC}^d$ , i.e. results of Section 2, hold for pictures of general prototype  $p : [n_1] \times \dots \times [n_d] \rightarrow \Sigma$ , without any restriction, i.e. for all  $n_i \geq 1$ ,  $i \in [d]$ . Moreover, our proofs also hold without change, except of course, the references to integer  $n$  which should be replaced by integer  $n_i$  according to the involved dimension  $i \in [d]$ .

#### $\text{NLIN}_{ca}^d$ and logical characterizations in coordinate encodings

The definition of linear time complexity of  $d$ -automata is not clear for input pictures whose domain has the general form  $[n_1] \times \dots \times [n_d]$  without restriction. However, linear time, i.e. time  $O(n)$ , makes sense when the  $n_i$  are of the same order  $\Theta(n)$ , i.e. the pictures are "well-balanced". This justifies the following definition.

**Definition 4.9.** A  $d$ -picture language  $L$  is well-balanced if there is some constant positive integer  $c$  such that the domain  $[n_1] \times \dots \times [n_d]$  of each picture  $p \in L$  fulfills the condition: for  $i \in [d]$ ,  $n \leq cn_i$ , where  $n = \max(n_1, \dots, n_d)$  is called the length of  $p$ ; we say that the picture  $p$  (resp. the picture language  $L$ ) is  $c$ -balanced.

It is straightforward to adapt our notion of linear time to well-balanced picture languages.

**Definition 4.10.** A  $c$ -balanced  $d$ -picture language  $L$  belongs to  $\text{NLIN}_{ca}^d$  if there exist a  $d$ -automaton  $\mathcal{A}$  and a linear function  $T(n) = c_1 n + c_2$  such that  $L$  is the set of  $c$ -balanced  $d$ -pictures  $p$  accepted by  $\mathcal{A}$  in time  $T(n)$ , where  $n$  is the length of  $p$ .

**Remark 4.11.** Those notions are justified by the following points:

- The "perimeter" of a  $c$ -balanced  $d$ -picture  $p$  of length  $n$  is  $\Theta(n)$  and its size (area, volume, etc., according to its dimension  $d$ ) is  $|p| = \Theta(n^d)$ .
- So, linear time means time linear in the perimeter of the  $d$ -picture  $p$  or, equivalently, in  $|p|^{1/d}$ .

We can easily extend all our results about square  $d$ -languages to well-balanced  $d$ -languages. In order to reduce the well-balanced case to the square case we need some new definitions.

**Definition 4.12.** To each  $d$ -picture  $p$  of length  $n$ , one associates its squared  $d$ -picture denoted  $p^\square$  of domain  $[n]^d$  obtained by putting the new special symbol  $\square$  in each additional cell. Formally,

$$p^\square(a) = \begin{cases} p(a) & \text{if } a \in \text{dom}(p); \\ \square & \text{otherwise.} \end{cases}$$

To any  $d$ -picture language  $L$  one associates its squared  $d$ -picture language

$$L^\square = \{p^\square : p \in L\}.$$

**Remark 4.13.** Let  $c$  be a constant. If  $p$  is a  $c$ -balanced  $d$ -picture, then the size of its squared picture  $p^\square$  is  $|p^\square| = \Theta(|p|)$

The following result is quite easy to prove.

**Lemma 4.14.** Let  $L$  be a  $c$ -balanced  $d$ -picture language. Then  $L \in \text{NLIN}_{ca}^d \Leftrightarrow L^\square \in \text{NLIN}_{ca}^d$ .

**Remark 4.15.** Notice that the nontrivial (right-to-left) implication of the Lemma 4.14 means that the  $d$ -automaton  $\mathcal{A}_1$  that recognizes  $L^\square$  in linear time can be simulated by some  $d$ -automaton  $\mathcal{A}_2$  that recognizes  $L$  in the same time (up to a constant factor) but with less space:  $\mathcal{A}_2$  only uses the cells of  $p$  instead of the cells of the squared picture  $p^\square$ . This is possible with the following trick: the fact that the  $i^{\text{th}}$  dimension  $n$  of  $p^\square$  is replaced by  $n_i \geq c^{-1}n$  allows to "fold"  $c$  times the picture  $p^\square$  along the  $i^{\text{th}}$  dimension  $n_i$  of  $p$ . All in all, each cell of  $p$  simulates (at most)  $c^d$  cells of  $p^\square$ . This is performed by taking for the set of states of  $\mathcal{A}_2$  the set  $\Gamma^{c^d}$  where  $\Gamma$  is the set of states of  $\mathcal{A}_1$ .

Now, let us compare any well-balanced  $d$ -picture language  $L$  and its squared language  $L^\square$  from a logical point of view. The domain of the coordinate representation of a  $d$ -picture  $p : [n_1] \times \dots \times [n_d]$  is naturally  $[n]$  where  $n$  is the length of  $p$ , i.e.  $n = \max(n_1, \dots, n_d)$ . So, we define the coordinate representation of  $p$  as

$$\text{coord}^d(p) = \text{coord}^d(p^\square).$$

Hence, as a trivial consequence of the definitions,

$$\text{coord}^d(L) = \text{coord}^d(L^\perp).$$

The previous definitions and remarks and Lemma 4.14 justify that the results of Sections 3 and 4 hold without change in the extended case of well-balanced  $d$ -picture languages.

## Conclusion

Notice that Theorem 3.7 that characterizes the *linear* time complexity class of nondeterministic *cellular automata* is very similar to the following result about time complexity  $O(n^d)$ , for any  $d \geq 1$ , of nondeterministic RAMs, by two of the present authors [GO04]:

$$\text{NTIME}_{\text{ram}}(n^d) = \text{ESOF}(\text{var } d) = \text{ESOF}(\forall^d, \text{arity } d).$$

The main difference is that this latter result involves the existential second-order logic *with functions* (ESOF) *instead of* or *in addition to relations* and holds in all kinds of structures without restriction: pictures, structures of any arity and any type, etc. It is also interesting and maybe surprising to notice that, in those results, the *time degree*  $d$  of a RAM computation plays the same role as the *dimension*  $d + 1$  of the time-space diagram of a linear time bounded computation for a  $d$ -dimensional cellular automaton.

Both results attest of the robustness of the time complexity classes  $\text{NTIME}_{\text{ram}}(n^d)$  and  $\text{NLIN}_{\text{ca}}^d$ . They stress the significance of the RAMs as a sequential model, and of the cellular automata as a parallel model.

Finally, if we continue the comparison, we have, for coordinate representation of  $d$ -languages and for  $d \geq 2$ , the (strict) inclusions:

$$\text{REC}^d \subsetneq \text{ESO}(\forall^d, \text{arity } d) \subsetneq \text{ESO}(\forall^{d+1}, \text{arity } d) \subseteq \text{NLIN}_{\text{ca}}^d$$

for cellular automata and ESO, which strongly contrast with the equalities:

$$\text{NTIME}_{\text{ram}}(n^d) = \text{ESOF}(\forall^1, \text{arity } 1) = \text{ESOF}(\forall^2, \text{arity } 1)$$

for RAMs and ESO [GO04, DGO04].

## References

- [Büc60] J. R. Büchi. Weak second-order arithmetic and finite automata. *Z. Math. Logik Grundlagen Math.*, 6:66–92, 1960.
- [CE12] B. Courcelle and J. Engelfriet. *Graph structure and monadic second-order logic, a language theoretic approach*. Cambridge University Press, to appear in 2012.
- [DG06] A. Durand and E. Grandjean. First-order queries on structures of bounded degree are computable with constant delay. *ACM Transactions on Computational Logic*, V:1–18, 2006.
- [DGO04] A. Durand, E. Grandjean, and F. Olive. New results on arity vs. number of variables. Research Report 20-2004, LIF - Université de Provence, 2004.
- [dR87] Michel de Rougemont. Second-order and inductive definability on finite structures. *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik*, 33:47–63, 1987.

- [EF95] H.-D. Ebbinghaus and J. Flum. *Finite Model Theory*. Springer-Verlag, 1995.
- [Fag74] R. Fagin. Generalized first-order spectra and polynomial-time recognizable sets. In R. M. Karp, editor, *Complexity of Computation, SIAM-AMS Proceedings*, pages 43–73, 1974.
- [Gai82] H. Gaifman. On local and nonlocal properties. In J. Stern, editor, *Logic Colloquium '81*, pages 105–135. North Holland, 1982.
- [GKL<sup>+</sup>07] Erich Grädel, Ph. G. Kolaitis, L. Libkin, M. Marx, J. Spencer, M. Y. Vardi, Y. Venema, and S. Weinstein. *Finite Model Theory and Its Applications*. Texts in Theoretical Computer Science. Springer, 2007.
- [GO04] E. Grandjean and F. Olive. Graph properties checkable in linear time in the number of vertices. *Journal of Computer and System Sciences*, 68:546–597, 2004.
- [GRST96] D. Giammarresi, A. Restivo, S. Seibert, and W. Thomas. Monadic second-order logic over rectangular pictures and recognizability by tiling systems. *Information and Computation*, 125(1):32 – 45, 1996.
- [Han65] W. Hanf. Model-theoretic methods in the study of elementary logic. In J. Addison, L. Henkin, and A. Tarski, editors, *The Theory of Models*, pages 132–145. North Holland, 1965.
- [Imm99] N. Immerman. *Descriptive Complexity*. Graduate Texts in Computer Science. Springer, 1999.
- [Lib04] Leonid Libkin. *Elements of Finite Model Theory*. Springer, 2004.
- [Mat98] Oliver Matz. One quantifier will do in existential monadic second-order logic over pictures. In *MFCS*, pages 751–759, 1998.
- [MT97] Oliver Matz and Wolfgang Thomas. The monadic quantifier alternation hierarchy over graphs is infinite. In *LICS*, pages 236–244, 1997.
- [Tho82] Wolfgang Thomas. Classifying regular events in symbolic logic. *J. Comput. Syst. Sci.*, 25(3):360–376, 1982.
- [Tur84] G. Turán. On the definability of properties of finite graphs. *Discrete Mathematics*, 49:291–302, 1984.
- [TW68] James W. Thatcher and Jesse B. Wright. Generalized finite automata theory with an application to a decision problem of Second-Order logic. *Mathematical Systems Theory*, 2(1):57–81, March 1968.

## Appendix A. Normalization of "guessed" relations

This appendix is dedicated to the proof of the normalization of  $\text{ESO}(\forall^d, \text{arity } d)$  on coordinate encodings of  $(d-1)$ -pictures, for  $d \geq 2$ . Our purpose is to rewrite  $\text{ESO}(\forall^d, \text{arity } d)$ -formulas into equivalent formulas that fulfill the "sorted" or "local" property mentioned in the proof of the "sufficient condition" of Proposition 3.8.

Before formalizing our notion of "sorted" formulas, let us detail what is its intended meaning. We want to deal with pixels of time-space diagram of the computation that are adjacents, that is, that are both connected and differ (by one) by at most one dimension. Two such pixels are represented by  $d$ -tuples of the form  $\mathbf{x}$  and  $\mathbf{x}^{(i)}$ , that is:

- their components are in the same order (elsewhere they could be disconnected);
- there is at most one occurrence of succ (elsewhere, they would differ of more than one dimension).

These requirements (and a little more) are formalized in the following definition.

**Definition Appendix A.1.** *Let  $k, d$  be two integers such that  $d \geq k \geq 1$ . A sentence over coordinate structures for  $k$ -pictures is in  $\text{ESO}(\forall^d, \text{arity } d, \text{sorted})$  if it is of the form  $\exists \mathbf{R} \forall \mathbf{x} \psi(\mathbf{x})$  where*

- $\mathbf{R}$  is a list of relation symbols of arity  $d$ ;
- $\psi$  is a quantifier-free formula whose list of first-order variables is  $\mathbf{x} = (x_1, \dots, x_d)$ ;
- each atom of  $\psi$  is of one of the following forms:
  - (i)  $Q_s(x_1, \dots, x_k)$ , for  $s \in \Sigma$ ,
  - (ii)  $R(\mathbf{x})$  or  $R(\mathbf{x}^{(i)})$  where  $R \in \mathbf{R}$ ,  $i \in [d]$ , and  $\mathbf{x}^{(i)}$  is the tuple  $\mathbf{x}$  where  $x_i$  is replaced by  $\text{succ}(x_i)$ ,
  - (iii)  $\min(x_i)$  or  $\max(x_i)$ , for  $i \in [d]$ .

We prove the normalization  $\text{ESO}(\forall^d, \text{arity } d) = \text{ESO}(\forall^d, \text{arity } d, \text{sorted})$  for  $(d-1)$ -pictures (i.e., for  $k = d-1$ ). In the present section, we deal with Condition (ii) of the above definition (see Proposition Appendix A.9). In Subsection Appendix A.1, we eliminate equalities and inequalities. At this point, we get a normalization of  $\text{ESO}(\forall^d, \text{arity } d)$  into the so-called "half-sorted logic", denoted by  $\text{ESO}(\forall^d, \text{arity } d, \text{half-sorted})$ . It remains to manage the input relation symbols; this is done in Section Appendix B, where we tackle Condition (i).

To lighten the presentation of the forthcoming results, we first introduce some notations about tuples and permutations.

**Definition Appendix A.2.** *Let  $n, d > 0$  and  $\mathbf{x} \in [n]^d$ .*

- We denote by  $[\mathbf{x}]_i$  the  $i^{\text{th}}$  component of  $\mathbf{x}$ . E.g.  $(5, 7, 2)_2 = 7$ .
- We say that  $\mathbf{x}$  is **non decreasing**, and we write  $\mathbf{x} \uparrow$ , when  $[\mathbf{x}]_1 \leq \dots \leq [\mathbf{x}]_d$ .
- $\mathcal{S}(d)$  stands for the set of permutations of  $\{1, \dots, d\}$ . Given  $\alpha_1, \dots, \alpha_d \in \{1, \dots, d\}$ , we denote by  $\alpha_1 \dots \alpha_d$  the permutation  $\alpha \in \mathcal{S}(d)$  that maps each  $i$  on  $\alpha_i$ . Conversely, for  $\alpha \in \mathcal{S}(d)$  we set  $\alpha_i := \alpha(i)$ . By  $\mathcal{T}(d)$  we denote the set of transpositions over  $\{1, \dots, d\}$ .
- If  $\alpha \in \mathcal{S}(d)$  and  $\mathbf{x} = (x_1, \dots, x_d)$ , we denote by  $\mathbf{x}_\alpha$  the  $d$ -tuple  $(x_{\alpha_1}, \dots, x_{\alpha_d})$ . It is less ambiguous to define  $\mathbf{x}_\alpha$  by the assertion:

for any  $i \in \{1, \dots, d\}$ ,  $[\mathbf{x}_\alpha]_i = [\mathbf{x}]_{\alpha(i)}$ .

Thus, if  $\beta$  also belongs to  $\mathcal{S}(d)$ , we get  $[(\mathbf{x}_\alpha)_\beta]_i = [\mathbf{x}_\alpha]_{\beta(i)} = [\mathbf{x}]_{\alpha\beta(i)}$ . Whence the identity:  $(\mathbf{x}_\alpha)_\beta = \mathbf{x}_{\alpha\beta}$ . In particular,  $(\mathbf{x}_\alpha)_{\alpha^{-1}} = \mathbf{x}$ .

- For  $\alpha \in \mathcal{S}(d)$  and  $n > 0$ , we set  $[\alpha] = \{\mathbf{x} \in [n]^d \text{ s.t. } \mathbf{x}_\alpha \uparrow\}$ . In particular, denoting by  $\text{id}$  the identity on  $\{1, \dots, d\}$ , we get  $\mathbf{x} \in [\text{id}]$  iff  $\mathbf{x} \uparrow$ . Therefore,  $\mathbf{x} \in [\alpha]$  iff  $\mathbf{x}_\alpha \in [\text{id}]$ . Clearly,  $[n]^d = \bigcup_{\alpha \in \mathcal{S}(d)} [\alpha]$ .
- For any  $i$  in  $\{1, \dots, d\}$ ,  $\mathbf{x}^{(i)}$  denotes the tuple obtained from  $\mathbf{x}$  by replacing its  $i^{\text{th}}$  component by its own successor. That is, if  $\mathbf{x} = (x_1, \dots, x_d)$  then

$$\mathbf{x}^{(i)} = (x_1, \dots, x_{i-1}, \text{succ}(x_i), x_{i+1}, \dots, x_d).$$

As previously, the arrangement of  $\mathbf{x}^{(i)}$  according to some permutation  $\alpha$  is denoted by  $(\mathbf{x}^{(i)})_\alpha$ .

**Example.** Consider  $\mathbf{x} = (5, 3, 7, 2)$  in  $[9]^4$  and  $\alpha = 4213$ ,  $\beta = 1432$  in  $\mathcal{S}(4)$ . Then  $\mathbf{x}_\alpha = (2, 3, 5, 7)$  is non-decreasing while  $\mathbf{x}_\beta = (5, 2, 7, 3)$  is not. Besides,  $\mathbf{x}^{(3)} = (5, 3, 8, 2)$  while  $(\mathbf{x}_\alpha)^{(3)} = (2, 3, 6, 7)$  and  $(\mathbf{x}^{(3)})_\alpha = (2, 3, 5, 8) = (\mathbf{x}_\alpha)^{(4)}$ .  $\triangleleft$

With these notations, the request described at the beginning of the section can be rephrased as follows: we want to normalize  $\text{ESO}(\forall^d, \text{arity } d)$ -formulas in such a way that each atomic subformula  $R(t_1(\mathbf{x}), \dots, t_p(\mathbf{x}))$  built with a guessed relation symbol  $R$  has either the form  $R(\mathbf{x})$  or the form  $R(\mathbf{x}^{(i)})$ .

**Fact Appendix A.3.** On  $\text{COORD}^{d-1}$ ,<sup>5</sup> any formula  $\Phi = \exists \mathbf{R} \forall \mathbf{x} \varphi(\mathbf{x}, \mathbf{R}, \sigma) \in \text{ESO}(\forall^d, \text{arity } d)$  of signature  $\sigma$  can be written in such a way that:

- (a) In each atomic subformula  $R(t_1, \dots, t_p)$  of  $\varphi$ ,  $\text{Var}(t_i) \cap \text{Var}(t_j) = \emptyset$  for every  $1 \leq i < j \leq p$ .
- (b) Each  $R$  in  $\mathbf{R}$  has arity  $d$  exactly.

**PROOF.** The proof of (a) is quite immediate. We illustrate it with an example: assume that  $\Phi$  involves the subformula  $R(\text{succ}^2 x_1, x_2, \text{succ } x_1, \text{succ}^3 x_2)$  for some  $R \in \sigma \cup \{\mathbf{R}\}$ . Then clearly,  $\Phi$  is equivalent, on picture-structures, to:

$$\exists \mathbf{R} \exists A \forall \mathbf{x} : \tilde{\varphi}(\mathbf{x}, \mathbf{R}, A, \sigma) \wedge (x_1 = \text{succ } x_3 \wedge x_4 = \text{succ}^3 x_2) \rightarrow (A(x_2, x_3) \leftrightarrow R(x_1, x_2, x_3, x_4))$$

where  $\tilde{\varphi}$  is obtained from  $\varphi$  by substituting the formula  $A(x_2, \text{succ}(x_1))$  to each occurrence of the atom  $R(\text{succ}^2 x_1, x_2, \text{succ } x_1, \text{succ}^3 x_2)$ .

In order to prove (b), assume for simplicity that  $\mathbf{R}$  reduces to the single relational symbol  $R$  of arity  $p < d$ . The idea is to add  $d - p$  dummy arguments to  $R$ . Clearly,  $\Phi$  is equivalent to the formula:

$$\exists R : \tilde{\varphi} \wedge \forall \mathbf{x} \bigwedge_{p < i \leq d} R(x_1, \dots, x_i, \dots, x_d) \leftrightarrow R(x_1, \dots, \text{succ}(x_i), \dots, x_d),$$

where  $\tilde{\varphi}$  is obtained from  $\varphi$  by replacing each atomic subformula  $R(t_1, \dots, t_p)$  by  $R(t_1, \dots, t_p, x_{i_1}, \dots, x_{i_{d-p}})$ . Here,  $x_{i_1}, \dots, x_{i_{d-p}}$  is the complete list of distinct variables among  $\mathbf{x}$  that do not occur in  $t_1, \dots, t_p$ .  $\square$

<sup>5</sup>We denote by  $\text{COORD}^{d-1}$  the sets of coordinate encodings of  $(d-1)$ -pictures.

**Remark Appendix A.4.** The proof of Fact [Appendix A.3](#) allows enhancing its statement: each atomic subformula of  $\tilde{\varphi}$  that involves an input symbol  $Q_a$ , for some  $a \in \Sigma$ , has the form  $Q_a(x_{\alpha_1}, \dots, x_{\alpha_{d-1}})$  where  $\alpha$  is an injection from  $\{1, \dots, d-1\}$  into  $\{1, \dots, d\}$ .

**Fact Appendix A.5.** On  $\text{COORD}^{d-1}$ , any formula  $\Phi = \exists \mathbf{R} \forall \mathbf{x} \varphi(\mathbf{x}, \mathbf{R}, \sigma) \in \text{ESO}(\forall^d, \text{arity } d)$  of signature  $\sigma$  can be written in such a way that each atomic subformula over  $\mathbf{R}$  of  $\varphi$  has one of the two forms:  $R(\mathbf{x}^{(i)})$  or  $R(\mathbf{x}_\pi)$ , where  $\pi \in \mathcal{S}(d)$ .

**PROOF.** We prove the result for  $d = 3$ . The general case is similar. Let  $\ell$  be the maximal value of an  $i \in \mathbb{N}$  such that  $\text{succ}^i(x)$  occurs in  $\Phi$ , for any  $x \in \mathbf{x}$ . For each  $R \in \mathbf{R}$ , we introduce new  $d$ -ary relation symbols  $R_{i,j,k}$  for every  $i, j, k \leq \ell$ . We want to force the following interpretations of the  $R_{i,j,k}$ 's:

$$R_{i,j,k}(u_1, u_2, u_3) = R(\text{succ}^i u_1, \text{succ}^j u_2, \text{succ}^k u_3).$$

This is done inductively, with the formulas:

- $\forall \mathbf{x} : R_{0,0,0}(x_1, x_2, x_3) \leftrightarrow R(x_1, x_2, x_3)$
- $\forall \mathbf{x} : \bigwedge_{i < \ell} \bigwedge_{j,k \leq \ell} (R_{i+1,j,k}(x_1, x_2, x_3) \leftrightarrow R_{i,j,k}(\text{succ}(x_1), x_2, x_3))$
- $\forall \mathbf{x} : \bigwedge_{j < \ell} \bigwedge_{i,k \leq \ell} (R_{i,j+1,k}(x_1, x_2, x_3) \leftrightarrow R_{i,j,k}(x_1, \text{succ}(x_2), x_3))$
- $\forall \mathbf{x} : \bigwedge_{k < \ell} \bigwedge_{i,j \leq \ell} (R_{i,j,k+1}(x_1, x_2, x_3) \leftrightarrow R_{i,j,k}(x_1, x_2, \text{succ}(x_3)))$

Factorizing the quantifications and using notations of Definition [Appendix A.2](#), the conjunction of these formulas can be written:

$$\forall \mathbf{x} \left\{ \begin{array}{l} R_{0,0,0}(\mathbf{x}) \leftrightarrow R(\mathbf{x}) \wedge \\ \bigwedge_{i < \ell} \bigwedge_{j,k \leq \ell} (R_{i+1,j,k}(\mathbf{x}) \leftrightarrow R_{i,j,k}(\mathbf{x}^{(1)})) \wedge \\ \bigwedge_{j < \ell} \bigwedge_{i,k \leq \ell} (R_{i,j+1,k}(\mathbf{x}) \leftrightarrow R_{i,j,k}(\mathbf{x}^{(2)})) \wedge \\ \bigwedge_{k < \ell} \bigwedge_{i,j \leq \ell} (R_{i,j,k+1}(\mathbf{x}) \leftrightarrow R_{i,j,k}(\mathbf{x}^{(3)})) \end{array} \right\}$$

Let us denote by  $\text{decomp}(R, (R_{i,j,k})_{i,j,k \leq \ell})$  this last formula. It clearly fulfills the condition of the statement. Now, consider the formula

$$\exists \mathbf{R} \exists ((R_{i,j,k})_{i,j,k \leq \ell})_{R \in \mathbf{R}} : \bigwedge_{R \in \mathbf{R}} \text{decomp}(R, (R_{i,j,k})_{i,j,k \leq \ell}) \wedge \forall \mathbf{x} \tilde{\varphi}, \quad (\text{A.1})$$

where  $\tilde{\varphi}$  is obtained from  $\varphi$  by the substitutions

$$R(\text{succ}^i x_{\alpha_1}, \text{succ}^j x_{\alpha_2}, \text{succ}^k x_{\alpha_3}) \rightsquigarrow R_{i,j,k}(x_{\alpha_1}, x_{\alpha_2}, x_{\alpha_3}).$$

Then, the formula (A.1) is equivalent to  $\Phi$  and also fits the requirements of Fact [Appendix A.5](#). It is the rewriting of  $\Phi$  announced.  $\square$



As a result of Fact [Appendix A.3](#), Remark [Appendix A.4](#) and Fact [Appendix A.5](#), each  $\text{ESO}(\forall^d, \text{arity } d)$ -formula of signature  $\sigma$  has a conjunctive normal form of the shape:

$$\Phi \equiv \exists \mathbf{R} \forall \mathbf{x} \bigwedge \bigvee \pm \left\{ \begin{array}{l} \min(\text{succ}^i(x)), \max(\text{succ}^i(x)), \\ \text{succ}^i(x) = \text{succ}^j(y), \\ Q_a(\mathbf{x}_i), R(\mathbf{x}_\beta), R(\mathbf{x}^{(i)}) \end{array} \right\}$$

Furthermore, the trick used in the proof of Fact [Appendix A.5](#) also allows writing atoms involving min, max or equalities, under the form  $\max(x)$ ,  $\min(x)$  and  $x = y$  for some  $x, y \in \mathbf{x}$ .

It remains to prove that we can get rid of the atomic formulas  $R(\mathbf{x}_\beta)$ , where  $\beta \neq \text{id}$ . This part is rather technical, so we provide some preliminary explanations before stating the logical framework which allows the normalization. In order to get rid of each literal of the form  $R(x_\beta)$ , we will divide the set  $R \subseteq [n]^d$  in  $d!$  relations  $R_\beta \subseteq [n]^d$ , each corresponding to a given permutation  $\beta$  of  $\{1, \dots, d\}$ .

**Definition [Appendix A.6](#).** For  $R \subseteq [n]^d$  and for each  $\alpha \in \mathcal{S}(d)$ , we define a  $d$ -ary relation  $R_\alpha$  on  $[n]$  by:  $R_\alpha = \{\mathbf{x} \in [\text{id}] \text{ s.t. } R(\mathbf{x}_{\alpha^{-1}})\}$ . Alternatively,  $R_\alpha$  can be defined by:  $R_\alpha = \{\mathbf{x}_\alpha : \mathbf{x} \in R \cap [\alpha]\}$ .

Thus, Definition [Appendix A.6](#) associates with each  $R \subseteq [n]^d$  a family  $(R_\alpha)_{\alpha \in \mathcal{S}(d)}$  of relations, each of which is entirely contained in the set  $[\text{id}]$ . This family is intended to represent  $R$  through its  $d!$  fragments according to the partition  $[n]^d = \bigcup_{\alpha \in \mathcal{S}(d)} [\alpha]$ . Namely, each  $R_\alpha$  encodes the fragment  $R \cap [\alpha]$  over  $[\text{id}]$ .

Actually,  $\bigcup_{\alpha \in \mathcal{S}(d)} [\alpha]$  is not really a partition, since the  $[\alpha]$ 's can overlap. Hence, Definition [Appendix A.6](#) induces some connexions between the relations  $R_\alpha$ : if some  $\mathbf{x}$  is both in  $[\alpha]$  and in  $[\beta]$ , or equivalently, if  $\mathbf{x}_\alpha = \mathbf{x}_\beta$ , then  $\mathbf{x} \in R \cap [\alpha]$  iff  $\mathbf{x} \in R \cap [\beta]$  and hence, by Definition [Appendix A.6](#):  $R_\alpha(\mathbf{x}_\alpha) = R_\beta(\mathbf{x}_\beta)$ . We will keep in mind :

$$\forall \alpha, \beta \in \mathcal{S}(d), \forall \mathbf{x} \in [n]^d : \mathbf{x}_\alpha = \mathbf{x}_\beta \Rightarrow R_\alpha(\mathbf{x}_\alpha) = R_\beta(\mathbf{x}_\beta). \quad (\text{A.2})$$

The following lemma states that condition (A.2) ensures that the  $R_\alpha$ 's issue from a single relation  $R$ , according to Definition [Appendix A.6](#). Besides, a new formulation of the condition is given in Item 3 of the lemma, that will better fit our syntactical restrictions.

**Lemma [Appendix A.7](#).** Let  $(R_\alpha)_{\alpha \in \mathcal{S}(d)}$  be a family of  $d$ -ary relations on  $[n]$  such that  $R_\alpha \subseteq [\text{id}]$  for each  $\alpha$ . The following are equivalent:

1.  $\exists R \subseteq [n]^d$  such that  $R_\alpha = \{\mathbf{x} \in [\text{id}] \text{ s.t. } R(\mathbf{x}_{\alpha^{-1}})\}$  for each  $\alpha \in \mathcal{S}(d)$  ;
2.  $\forall \alpha, \beta \in \mathcal{S}(d), \forall \mathbf{x} \in [n]^d : \mathbf{x}_\alpha = \mathbf{x}_\beta \Rightarrow R_\alpha(\mathbf{x}_\alpha) = R_\beta(\mathbf{x}_\beta)$  ;
3.  $\forall \alpha \in \mathcal{S}(d), \forall \tau \in \mathcal{T}(d), \forall \mathbf{x} \in [n]^d : \mathbf{x} = \mathbf{x}_\tau \Rightarrow R_\alpha(\mathbf{x}) = R_{\alpha\tau}(\mathbf{x})$ .  
(Recall  $\mathcal{T}(d)$  denotes the set of transpositions over  $\{1, \dots, d\}$ .)

PROOF.  $1 \Rightarrow 2$ : See Equation (A.2).

$2 \Rightarrow 1$ : For  $(R_\alpha)_{\alpha \in \mathcal{S}(d)}$  fulfilling 2, consider the relation  $R \subseteq [n]^d$  defined by:

$$R(\mathbf{x}) \text{ iff } R_\alpha(\mathbf{x}_\alpha) \text{ for some } \alpha \text{ such that } \mathbf{x}_\alpha \uparrow. \quad (\text{A.3})$$

This definition is well formed, since for any  $\alpha, \beta \in \mathcal{S}(d)$  and any  $\mathbf{x} \in [n]^d$  such that both  $\mathbf{x}_\alpha \uparrow$  and  $\mathbf{x}_\beta \uparrow$  hold, we have  $\mathbf{x}_\alpha = \mathbf{x}_\beta$  and thus, by 2,  $R_\alpha(\mathbf{x}_\alpha) = R_\beta(\mathbf{x}_\beta)$ . Now, let  $\alpha \in \mathcal{S}(d)$ . For any  $\mathbf{x} \in [\text{id}]$  we have  $(\mathbf{x}_{\alpha^{-1}})_\alpha \uparrow$  (since  $(\mathbf{x}_{\alpha^{-1}})_\alpha = \mathbf{x}$ ) and hence, by (A.3),  $R_\alpha(\mathbf{x}) = R_\alpha((\mathbf{x}_{\alpha^{-1}})_\alpha) = R(\mathbf{x}_{\alpha^{-1}})$ . Besides,  $R_\alpha(\mathbf{x}) = 0$  for any  $\mathbf{x} \notin [\text{id}]$ , since  $R_\alpha \subseteq [\text{id}]$ . Thus  $R_\alpha$  is obtained from  $R$  as required in 1.

2  $\Rightarrow$  3: Let  $\alpha \in \mathcal{S}(d)$ ,  $\tau \in \mathcal{T}(d)$  and  $\mathbf{x} \in [n]^d$  such that  $\mathbf{x} = \mathbf{x}_\tau$ . Set  $\mathbf{y} = \mathbf{x}_{\alpha^{-1}}$ . Then,  $\mathbf{y}_\alpha = \mathbf{x} = \mathbf{x}_\tau = (\mathbf{y}_\alpha)_\tau = \mathbf{y}_{\alpha\tau}$ . Therefore we get by 2:  $R_\alpha(\mathbf{y}_\alpha) = R_{\alpha\tau}(\mathbf{y}_{\alpha\tau})$ , and hence:  $R_\alpha(\mathbf{x}) = R_{\alpha\tau}(\mathbf{x})$ .

3  $\Rightarrow$  2: Let  $\alpha, \beta \in \mathcal{S}(d)$  and  $\mathbf{x} \in [n]^d$  such that  $\mathbf{x}_\alpha = \mathbf{x}_\beta$ . For  $\mathbf{y} = \mathbf{x}_\alpha$ , the equality  $\mathbf{x}_\alpha = \mathbf{x}_\beta$  can be written  $\mathbf{y} = \mathbf{y}_{\alpha^{-1}\beta}$ . It means that the permutation  $\alpha^{-1}\beta$  exchanges integers that index equal components of  $\mathbf{y}$ . It is easily seen that this property can be required for each transposition occurring in a decomposition of  $\alpha^{-1}\beta$  on  $\mathcal{T}(d)$ . That is, there exist some transpositions  $\tau_1, \dots, \tau_k \in \mathcal{T}(d)$  such that  $\alpha^{-1}\beta = \tau_1 \dots \tau_k$  and  $\mathbf{y} = \mathbf{y}_{\tau_1} = \mathbf{y}_{\tau_1\tau_2} = \dots = \mathbf{y}_{\tau_1 \dots \tau_k}$ . Then, applying 3 to these successive tuples, we get:  $R_\alpha(\mathbf{y}) = R_{\alpha\tau_1}(\mathbf{y}_{\tau_1}) = R_{\alpha\tau_1\tau_2}(\mathbf{y}_{\tau_1\tau_2}) = \dots = R_{\alpha\tau_1 \dots \tau_k}(\mathbf{y}_{\tau_1 \dots \tau_k})$ . Hence  $R_\alpha(\mathbf{y}) = R_\beta(\mathbf{y}_{\alpha^{-1}\beta})$ , that is  $R_\alpha(\mathbf{x}_\alpha) = R_\beta(\mathbf{x}_\beta)$ , as required.  $\square$

**Lemma Appendix A.8.** Let  $R$  and  $(R_\alpha)_{\alpha \in \mathcal{S}(d)}$  be defined as in Definition Appendix A.6. Let  $\alpha, \beta \in \mathcal{S}(d)$  and  $i \in \{1, \dots, d\}$ . For any  $\mathbf{x} \in [\alpha]$ :

1.  $R(\mathbf{x}_\beta)$  is equivalent to  $R_{\beta^{-1}\alpha}(\mathbf{x}_\alpha)$ .
2.  $R(\mathbf{x}^{(i)})$  is equivalent to:

$$\left( x_i = x_{\alpha_d} \wedge R_{\alpha\tau_{\alpha^{-1}(i),d}}((x_\alpha)^{(d)}) \right) \vee \bigvee_{i \leq k < d} \left( x_i = x_{\alpha_k} < x_{\alpha_{k+1}} \wedge R_{\alpha\tau_{\alpha^{-1}(i),k}}((x_\alpha)^{(k)}) \right).$$

PROOF. 1. We have seen that  $R(\mathbf{x})$  holds iff  $R_\gamma(\mathbf{x}_\gamma)$  holds for any  $\gamma \in \mathcal{S}(d)$  such that  $\mathbf{x}_\gamma \uparrow$ . Thus we get, for a given  $\beta \in \mathcal{S}(d)$ :  $R(\mathbf{x}_\beta)$  iff  $R_\gamma((\mathbf{x}_\beta)_\gamma)$  holds for any  $\gamma \in \mathcal{S}(d)$  such that  $(\mathbf{x}_\beta)_\gamma \uparrow$ . That is, since  $(\mathbf{x}_\beta)_\gamma = \mathbf{x}_{\beta\gamma}$ :

$$R(\mathbf{x}_\beta) \text{ iff } R_\gamma(\mathbf{x}_{\beta\gamma}) \text{ holds for any } \gamma \in \mathcal{S}(d) \text{ such that } \mathbf{x}_{\beta\gamma} \uparrow. \quad (\text{A.4})$$

In particular,  $\mathbf{x}_{\beta\gamma} \uparrow$  iff  $\mathbf{x}_{\beta\gamma} = \mathbf{x}_\alpha$ , since  $\mathbf{x} \in [\alpha]$ . Thus,  $\gamma = \beta^{-1}\alpha$  is one of the permutations such that  $\mathbf{x}_{\beta\gamma} \uparrow$ . Thus, replacing  $\gamma$  by  $\beta^{-1}\alpha$  in Equation (A.4), we get the sought result.

2. From  $\mathbf{x}^{(i)} = (x_1, \dots, \text{succ}(x_i), \dots, x_d)$  we get:

$$(\mathbf{x}^{(i)})_\alpha = (x_{\alpha_1}, \dots, x_{\alpha_{j-1}}, \text{succ}(x_i), x_{\alpha_{j+1}}, \dots, x_{\alpha_d})$$

where  $j = \alpha^{-1}(i)$ . Since  $x_{\alpha_1} \leq \dots \leq x_{\alpha_d}$ , the above tuple  $(\mathbf{x}^{(i)})_\alpha$  is almost increasingly ordered. More precisely, there exists  $k \in \{1, \dots, d\}$  such that:

$$x_{\alpha_1} \leq \dots \leq x_{\alpha_{j-1}} \leq x_i = x_{\alpha_{j+1}} = \dots = x_{\alpha_k} \leq x_{\alpha_{k+1}} \leq \dots \leq x_{\alpha_d},$$

where  $j = \alpha^{-1}(i)$ . Clearly, the largest such  $k$  is characterized by:  $(k = d)$  or  $(k < d \text{ and } x_i = x_{\alpha_k} < x_{\alpha_{k+1}})$ . Or equivalently, by:

$$(x_i = x_{\alpha_d}) \text{ or } (k < d \text{ and } x_i = x_{\alpha_k} < x_{\alpha_{k+1}}). \quad (\text{A.5})$$

If we denote by  $\tau_{j,k}$  the transposition over  $\{1, \dots, d\}$  which permutes  $j$  and  $k$ , the definition of  $k$  yields that the tuple

$$(\mathbf{x}^{(i)})_{\alpha\tau_{j,k}} = (x_{\alpha_1}, \dots, x_{\alpha_{j-1}}, \boxed{x_{\alpha_k}}, x_{\alpha_{j+1}}, \dots, x_{\alpha_{k-1}}, \boxed{\text{succ}(x_i)}, x_{\alpha_{k+1}}, \dots, x_{\alpha_d})$$

is non decreasing. Hence,  $R(\mathbf{x}^{(i)}) = R_{\alpha\tau_{j,k}}((\mathbf{x}^{(i)})_{\alpha\tau_{j,k}})$ . Besides, since  $x_{\alpha_k} = x_i$ , the tuple  $(\mathbf{x}^{(i)})_{\alpha\tau_{j,k}}$  above can also be written:

$$(\mathbf{x}^{(i)})_{\alpha\tau_{j,k}} = (x_{\alpha_1}, \dots, x_{\alpha_{j-1}}, \boxed{x_i}, x_{\alpha_{j+1}}, \dots, x_{\alpha_{k-1}}, \boxed{\text{succ}(x_{\alpha_k})}, x_{\alpha_{k+1}}, \dots, x_{\alpha_d}).$$

That is:  $(x^{(i)})_{\alpha\tau_{j,k}} = (x_\alpha)^{(k)}$ . Therefore:  $R(\mathbf{x}^{(i)}) = R_{\alpha\tau_{j,k}}((x_\alpha)^{(k)})$ . Reminding that  $j = \alpha^{-1}(i)$ , we can finally state: there exists a sole  $k \in \{i, \dots, d\}$  defined by (A.5), and for this  $k$  we have:  $R(\mathbf{x}^{(i)}) = R_{\alpha\tau_{\alpha^{-1}(i),k}}((x_\alpha)^{(k)})$ . The conclusion easily proceeds.  $\square$

**Proposition Appendix A.9.** For  $d > 1$ ,  $\text{ESO}(\forall^d, \text{arity } d) \subseteq \text{ESO}(\forall^d, \text{arity } d, \text{half-sorted})$  on  $\text{COORD}^{d-1}$ .

**PROOF.** To simplify, assume we want to translate in  $\text{ESO}(\forall^d, \text{arity } d, \text{half-sorted})$  some  $\text{ESO}(\forall^d, \text{arity } d)$ -formula of the very simple shape:  $\Phi \equiv \exists R \forall \mathbf{x} \varphi(\mathbf{x}, R)$ , where  $R$  is a (single)  $d$ -ary relation symbol,  $\mathbf{x}$  is a  $d$ -tuple of first-order variables, and  $\varphi$  is a quantifier-free formula. Since the sets  $[\alpha]$ ,  $\alpha \in \mathcal{S}(d)$ , cover the domain  $[n]$ , we obtain an equivalent rewriting of  $\Phi$  with the following artificial relativization:  $\Phi \equiv \exists R \forall \mathbf{x} \bigwedge_{\alpha \in \mathcal{S}(d)} (\mathbf{x} \in [\alpha] \rightarrow \varphi)$ . Furthermore, all atomic subformulas of  $\varphi$  built on  $R$  can be assumed of the form  $R(\mathbf{x}_\beta)$  or  $R(\mathbf{x}^{(i)})$ , thanks to Fact Appendix A.5.

To get rid of these literals, we substitute to  $R$  a tuple of relations  $(R_\alpha)_{\alpha \in \mathcal{S}(d)}$  that encode  $R$  on the sets  $[\alpha]$ . Recall we proved in Lemma Appendix A.7 that this substitution is legal as soon as  $R_\alpha \subseteq [\text{id}]$  and  $R_\alpha(\mathbf{x}) = R_{\alpha\tau}(\mathbf{x})$  for all  $\alpha \in \mathcal{S}(d)$ ,  $\tau \in \mathcal{T}(d)$  and every  $\mathbf{x} \in [n]^d$  such that  $\mathbf{x}_\tau = \mathbf{x}$ . Then, Lemma Appendix A.8 gives the translation of  $R$ -atomic formulas into formulas expressed in term of the  $R_\alpha$ 's. All in all, we get the equivalence of the initial formula  $\Phi$  to the following:

$$\exists (R_\alpha)_{\alpha \in \mathcal{S}(d)} \left\{ \begin{array}{l} \forall \mathbf{x} \bigwedge_{\alpha \in \mathcal{S}(d)} (R_\alpha(\mathbf{x}) \rightarrow \mathbf{x} \in [\text{id}]) \wedge \\ \forall \mathbf{x} \bigwedge_{\alpha \in \mathcal{S}(d)} \bigwedge_{\tau \in \mathcal{T}(d)} (\mathbf{x}_\tau = \mathbf{x} \rightarrow (R_\alpha(\mathbf{x}) \leftrightarrow R_{\alpha\tau}(\mathbf{x}))) \wedge \\ \forall \mathbf{x} \bigwedge_{\alpha \in \mathcal{S}(d)} (\mathbf{x} \in [\alpha] \rightarrow \varphi_\alpha(\mathbf{x}, (R_\gamma)_{\gamma \in \mathcal{S}(d)})) \end{array} \right\} \quad (\text{A.6})$$

where each  $\varphi_\alpha$  is obtained from  $\varphi$  by the substitutions:

- $R(\mathbf{x}_\beta) \rightsquigarrow R_{\beta^{-1}\alpha}(x_\alpha)$
- $R(\mathbf{x}^{(i)}) \rightsquigarrow \left\{ \begin{array}{l} \left( x_i = x_{\alpha_d} \wedge R_{\alpha\tau_{\alpha^{-1}(i),d}}((x_\alpha)^{(d)}) \right) \vee \\ \bigvee_{i \leq k < d} \left( x_i = x_{\alpha_k} < x_{\alpha_{k+1}} \wedge R_{\alpha\tau_{\alpha^{-1}(i),k}}((x_\alpha)^{(k)}) \right) \end{array} \right\}$

The first two conjuncts of (A.6) ensure that the family  $(R_\alpha)_{\alpha \in \mathcal{S}(d)}$  encodes a relation  $R$  (see Lemma Appendix A.7); the third interprets assertions of the form  $R(\mathbf{x}_\beta)$  and  $R(\mathbf{x}^{(i)})$  according to the modalities described in Lemma Appendix A.8. Because of permutability of the conjunction and the universal quantifier, this third conjunct can be rewritten:

$$\bigwedge_{\alpha \in \mathcal{S}(d)} \forall \mathbf{x} : \mathbf{x} \in [\alpha] \rightarrow \varphi_\alpha(\mathbf{x}, (R_\gamma)_{\gamma \in \mathcal{S}(d)}) \quad (\text{A.7})$$

For a fixed conjunct in (A.7), i.e. for a fixed  $\alpha$ , all atomic subformulas of  $\varphi_\alpha$  built on the  $R_\gamma$ 's have the form  $R_\gamma(\mathbf{x}_\alpha)$  or  $R_\gamma((\mathbf{x}_\alpha)^{(i)})$  for some  $\gamma \in \mathcal{S}(d)$  and some  $i \in \{1, \dots, d\}$ . Hence, the substitution of variables  $\mathbf{x}/\mathbf{x}_\alpha$  allows to write such a conjunct as:  $\forall \mathbf{x} : \mathbf{x} \in [\text{id}] \rightarrow \tilde{\varphi}_\alpha$  where  $\tilde{\varphi}_\alpha \equiv \varphi_\alpha(\mathbf{x}/\mathbf{x}_\alpha)$  only involves  $(R_\gamma)$ -subformulas

of the form  $R_\gamma(\mathbf{x})$  or  $R_\gamma(\mathbf{x}^{(k)})$  for some  $\gamma \in S(d)$  and  $k \in \{1, \dots, d\}$ . Finally, the initial formula  $\Phi$  is proved equivalent to:

$$\exists(R_\alpha)_{\alpha \in S(d)} \left\{ \begin{array}{l} \forall \mathbf{x} \bigwedge_{\alpha \in S(d)} (R_\alpha(\mathbf{x}) \rightarrow \mathbf{x} \in [\text{id}]) \wedge \\ \forall \mathbf{x} \bigwedge_{\alpha \in S(d)} \bigwedge_{\tau \in T(d)} (\mathbf{x}_\tau = \mathbf{x} \rightarrow (R_\alpha(\mathbf{x}) \leftrightarrow R_{\alpha\tau}(\mathbf{x}))) \wedge \\ \forall \mathbf{x} \bigwedge_{\alpha \in S(d)} (\mathbf{x} \in [\text{id}] \rightarrow \tilde{\varphi}_\alpha(\mathbf{x}, (R_\gamma)_{\gamma \in S(d)})) \end{array} \right\}$$

that fulfills the requirement of Proposition [Appendix A.9](#).  $\square$

#### Appendix A.1. Getting rid of arithmetic

Does mean introducing new second order variables of arity 2, we can assume that our formula  $\Phi \in \text{ESO}^\sigma(\mathcal{V}^d, \text{arity } d)$  – where each relation symbol of arity  $d \leq 2$  occurs in normalised (i.e. sorted) form – involves no comparison (equality, inequality) relation. We obtain that in two successive steps.

First, if  $\Phi$  involves inequalities  $<$  and  $>$ , then it is equivalent to the following formula  $\Phi'$  without inequalities (but with equalities) and two new binary relation symbols  $<$  and  $>$ :

$$\Phi' \equiv \exists < \exists >: \tilde{\Phi} \wedge \forall x_1, x_2 \left\{ \begin{array}{ll} (x_1 = x_2 \vee x_1 < x_2) \rightarrow (\neg \max(x_2) \rightarrow x_1 < \text{succ}(x_2)) & \wedge \\ (x_1 = x_2 \vee x_1 > x_2) \rightarrow (\neg \max(x_1) \rightarrow \text{succ}(x_1) > x_2) & \wedge \\ x_1 < x_2 \rightarrow (\neg(x_1 > x_2) \wedge \neg(x_1 = x_2)) & \wedge \\ x_1 > x_2 \rightarrow \neg(x_1 = x_2) & \end{array} \right\}$$

where  $\tilde{\Phi}$  is obtained from  $\Phi$  by the substitutions  $u < v \rightsquigarrow u < v$  (resp.  $u > v \rightsquigarrow u > v$ ). This is justified as follows: the first two conjuncts of the subformula  $\forall x_1 \forall x_2 \{\dots\}$  express that  $x_1 < x_2 \Rightarrow x_1 < x_2$  (resp.  $x_1 > x_2 \Rightarrow x_1 > x_2$ ). The third and fourth conjuncts express that the three relations  $<, >$  and  $=$  are totally disjoint. That implies that  $<$  and  $>$  have their exact meaning. Notice also that the occurrences of  $<$  and  $>$  preserve the sorted property.

Secondly, if  $\Phi$  involves equalities (without any inequality), it is equivalent to the following formula  $\Phi'$ , written without the symbol  $=$  but with the new binary symbol  $\approx$ :

$$\Phi' \equiv \exists \approx: \tilde{\Phi} \wedge \forall x_1, x_2 \Psi$$

where  $\tilde{\Phi}$  is obtained from  $\Phi$  by replacing each equality  $u = v$  by  $u \approx v$ , and  $\Psi$  is the conjunction of the following formulas:

- $\min(x_1) \rightarrow (\min(x_2) \leftrightarrow x_1 \approx x_2)$ ;
- $\min(x_2) \rightarrow (\min(x_1) \leftrightarrow x_1 \approx x_2)$ ;
- $(\neg \max(x_1) \wedge \neg \max(x_2)) \rightarrow (x_1 \approx x_2 \leftrightarrow \text{succ}(x_1) \approx \text{succ}(x_2))$ .

Notice that this transformation preserves sorted property. The results obtained until now can be recapitulated as follows:

On  $(d-1)$ -pictures, each  $\text{ESO}^\sigma(\forall^d, \text{arity } d)$ -formula can be written under the form:

$$\Phi \equiv \exists \mathbf{R} \forall \mathbf{x} \bigwedge \bigvee \pm \left\{ \begin{array}{l} \min(x), \max(x), \\ Q(\mathbf{x}_\iota), R(\mathbf{x}), R(\mathbf{x}^{(i)}) \end{array} \right\} \quad (\text{A.8})$$

Here,  $\mathbf{R}$  (resp.  $\mathbf{Q}$ ) is a list of relation symbols of arity  $d$  (resp.  $d-1$ ),  $x \in \mathbf{x}$ ,  $Q \in \cup(Q_\alpha)_{\alpha \in \Sigma}$ ,  $R \in \mathbf{R}$ ,  $\iota \in I(d)$ ,  $i \in \{1, \dots, d\}$ .

## Appendix B. Normalization of input relations

Let's come back to the normalization of the formula  $\Phi$  of Equation (A.8). For simplicity, assume that it has the restricted form  $\Phi \equiv \exists \mathbf{R} \forall \mathbf{x} \varphi(\mathbf{x}, \mathbf{R}, \mathbf{Q}, \sigma)$ , with only one relation symbol  $Q$  of arity  $d-1$ .

We aim at defining a tuple of relations  $(Q_\alpha)_{\alpha \in I(d)}$ , in such a way that  $Q_\alpha(\mathbf{x}) = Q(\mathbf{x}_\alpha)$  for each  $\mathbf{x}$ . Clearly, such relations will allow to write  $\Phi$  under the desired form, by replacing each subformula  $Q(\mathbf{x}_\alpha)$  by the sorted translation:  $Q_\alpha(\mathbf{x})$ . The difficulty is to express this definition with our syntactical restrictions, that is, without involving any  $\mathbf{x}_\alpha$  with  $\alpha \neq \text{id}$ .

Notice that the strategy used in Section Appendix A to "sort" atomic subformulas  $R(\mathbf{x}_\alpha)$  build on any *existentially quantified*  $d$ -ary relation  $R$  is no more available, since it means suppressing  $R$  in favour of some new existentially quantified relations. Of course, we can't operate like this with the *input* relation  $Q$ . To give an hint of the method used in this section, let us consider an easy example.

*An easy case..* Consider the case where  $d = 2$ . We deal with two first-order variables  $x$  and  $y$  and we only accept atoms of the form  $Q(x)$ ,  $R(x, y)$ ,  $R(\text{succ}(x), y)$  and  $R(x, \text{succ}(y))$  for any input unary relation  $Q$  and any guessed binary relation  $R$ . How can we tackle occurrences of some atom  $Q(y)$  in the formula? A natural idea is to define a new binary relation  $Q_2$  in such a way that  $Q_2(x, y) = Q(y)$  holds for any  $x, y$ . (We denote it  $Q_2$  to refer both to  $Q$  and to the projection of  $(x, y)$  on its *second* component.) Hence, we set:

$$Q_2 = \{(x, y) : Q(y)\}.$$

Thus, any atom  $Q(y)$  could be replaced by the sorted atom  $Q_2(x, y)$ . But the logical definition of  $Q_2$  with our syntactical constraints compels to introduce an additive binary relation  $T$  that will be used as a buffer to *transport* the information  $Q(y)$  into the expression  $Q_2(x, y)$ . We set

$$T = \{(x, y) : Q(x+y)\}.$$

Clearly,  $T$  is inductively defined from  $Q$  by the assertions  $T(x, 0) = Q(x)$  and  $T(x+1, y) = T(x, y+1)$ . Besides,  $Q_2$  is defined from  $T$  by  $Q_2(0, y) = T(0, y)$  and  $Q_2(x, y) = Q_2(x+1, y)$ . All these assertions can be rephrased in our logical framework, with the following formulas:

$$\begin{aligned} & \forall x, y \left\{ \begin{array}{ll} \min(y) \rightarrow (T(x, y) \leftrightarrow Q(x)) \wedge \\ (\neg \max(x) \wedge \neg \max(y)) \rightarrow (T(\text{succ}(x), y) \leftrightarrow T(x, \text{succ}(y))) \end{array} \right\} \\ & \forall x, y \left\{ \begin{array}{l} \min(x) \rightarrow (Q_2(x, y) \leftrightarrow T(x, y)) \wedge \\ Q_2(x, y) \leftrightarrow Q_2(\text{succ}(x), y) \end{array} \right\} \end{aligned}$$

Now, it remains to insert this defining formulas in the initial formula  $\Phi$  to be normalized, and to replace each occurrence of  $Q(y)$  by  $Q_2(x, y)$ . Of course, such a construction has to be carried on for each input unary relation  $Q$ .

*The general case..* Given  $i, j \in \{1, \dots, d\}$  and  $\alpha \in \mathcal{S}(d)$ , we denote by  $(ij)$  the transposition that exchanges  $i$  and  $j$ , and by  $\alpha(ij)$  the composition of  $\alpha$  and  $(ij)$ . It is well-known that each permutation  $\alpha$  can be written as a product of transpositions,  $\alpha = (u_1 v_1)(u_2 v_2) \dots (u_p v_p)$ . It is easily seen that this product can be chosen in such a way that  $v_i = u_{i+1}$  for any  $i$ . This is because if some sequence  $(ab)(cd)$  with  $b \neq c$  occurs, it can be replaced by  $(ca)(ab)(bc)(cd)$ , and a well chosen iteration of such rewritings yields the desired decomposition. This can be further refined, by fixing at  $d$  one element of the first transposition involved in the decomposition and by prohibiting useless sequence as  $(ab)(ba)$ . Finally, each  $\alpha \in \mathcal{S}(d)$  can be written

$$\alpha = (du_1)(u_1 u_2) \dots (u_{k-2} u_{k-1})(u_{k-1} u_k), \quad (\text{B.1})$$

where  $u_i, u_{i+1}$  and  $u_{i+2}$  are pairwise distinct elements of  $\{1, \dots, d\}$  for any  $i$ . We call *alternated factorization of  $\alpha$*  such a decomposition.

A permutation  $\alpha$  admits of several alternated factorizations, and we want to single out one of them for each  $\alpha$ , in order to allow an inductive reasoning build on this particular decomposition. There is no canonical way to perform this task. In the following lemma, we roughly describe one possible choice, that implicitly refers to the graph  $\mathcal{G}_d$  on domain  $\mathcal{S}(d)$  whose edges correspond to those pairs of permutations  $(\alpha, \beta)$  such that  $\beta = \alpha(ij)$  for some  $i, j \in \{1, \dots, d\}$ .

**Lemma Appendix B.1.** *There exists an oriented tree  $\mathcal{T}_d$  covering  $\mathcal{S}(d)$  which is rooted at  $\text{id}$  and such that each  $\mathcal{T}_d$ -path starting at  $\text{id}$ , say  $\text{id}\alpha_1 \dots \alpha_p$ , corresponds to an alternated factorization of  $\alpha_p$ .*

**PROOF.** Trees  $\mathcal{T}_d$  for  $d \geq 2$  are defined inductively. For  $d = 2$ , there is a unique such tree:  $(12) \rightarrow (21) = (12)(21)$ . So, assume we are given  $\mathcal{T}_{d-1}$  and carry on the construction of  $\mathcal{T}_d$  as follows:

- (a) First, view each permutation  $\alpha = \alpha_1 \dots \alpha_{d-1} \in \mathcal{S}(d-1)$  as a permutation of  $\{2, \dots, d\}$  by renaming  $\alpha_i$  as  $\alpha_i + 1$ . That is, replace  $\alpha = \alpha_1 \dots \alpha_{d-1}$  by  $\alpha^+ = (\alpha_1 + 1) \dots (\alpha_{d-1} + 1)$ .
- (b) Then, replace each such  $\alpha^+$  by  $\beta = \beta_1 \dots \beta_d \in \mathcal{S}(d)$  defined by:  $\beta_1 = 1$  and  $\beta_i = [\alpha^+]_{i-1} = \alpha_{i-1} + 1$  for  $i > 1$ . Thus,  $\mathcal{T}_{d-1}$  now covers the set of permutations  $\beta \in \mathcal{S}(d)$  that fulfill  $[\beta]_1 = 1$ .
- (c) For each node  $\beta$  thus obtained, create a new node labelled by the composition of  $\beta$  with the transposition  $(1d)$  – that is by the permutation  $\beta(1d)$  – and create an edge  $\beta \rightarrow \beta(1d)$ .
- (d) Finally, link each such node  $\beta(1d)$  to  $d-2$  new nodes  $\beta(1d)(di)$ , for  $i = 2, \dots, d-1$ .

In Fig B.1, we display the steps of the construction of  $\mathcal{T}_4$  from  $\mathcal{T}_2$ . Letters (a), ..., (d) in the figure refer to the above items. The correction of the method on this example is clear. We leave it to the reader to verify that it generalises to any  $d$ .  $\square$

This lemma allows us to choose, for each  $\alpha \in \mathcal{S}(d)$ , *one* alternated factorization of  $\alpha$ : it is the decomposition  $(di_1)(i_1 i_2) \dots (i_{k-1} i_k)$  corresponding to the unique path from  $\text{id}$  to  $\alpha$  in  $\mathcal{T}_d$ . We denote by  $\text{id}.di_1.i_1 i_2. \dots .i_{k-1} i_k$  this particular factorization. And when this path until  $\alpha$  can be continued in  $\mathcal{T}_d$  to some permutation  $\alpha(i_k i_{k+1})$ , we denote by  $\alpha.i_k i_{k+1}$  this last permutation. For instance, in the example displayed in Fig B.1, we can write  $2143 = 4123.13$  and  $3124 = 4123.14$  while  $4321 = 4123(24)$  *cannot be written*  $4123.24$ . Notice furthermore that the integers  $i_k$  and  $i_{k+1}$  are ordered in the notation  $\alpha.i_k i_{k+1}$  (unlike in the notation  $\alpha(i_k i_{k+1})$ ): we place in first position the integer  $i_k$  involved in the last transposition leading to  $\alpha$  (with  $i_k = d$  if  $\alpha = \text{id}$ ). All in all, the reader is invited to keep in mind the numerous presuppositions attached to the notation  $\alpha.uv$ : the statement  $\beta = \alpha.uv$  means:

- $\beta = \alpha(uv)$  ;

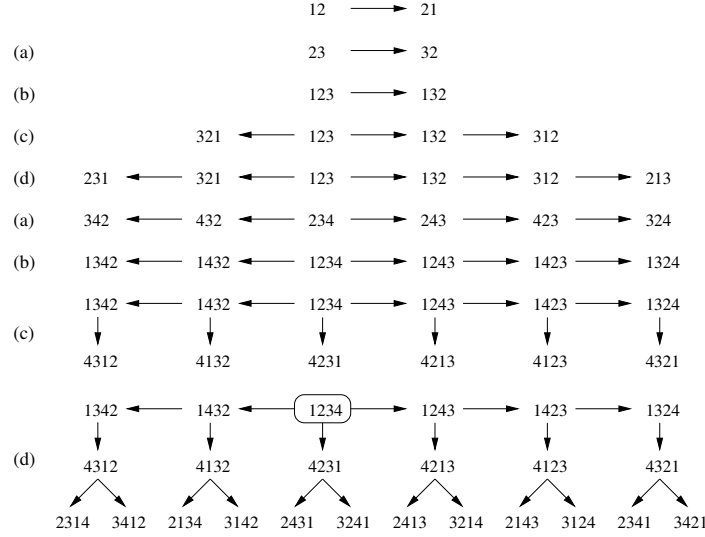


Figure B.1: Construction of  $\mathcal{T}_4$  from  $\mathcal{T}_2$ . The result is an oriented tree rooted at  $\text{id}$ , spanning  $\mathcal{S}(4)$ , whose all paths from  $\text{id}$  are alternated.

- $(\alpha, \beta)$  is an edge of  $\mathcal{T}_d$  ;
- either  $\alpha = \text{id}$  and  $u = d$ , or  $\alpha = \gamma.tu$  for some  $\gamma \in \mathcal{S}(d)$  and some  $t \neq u$  in  $\{1, \dots, d\}$ .

Let us now come to a straightforward remark connecting  $d$ -tuples to alternated factorizations. First recall that for  $\mathbf{x} = (x_1, \dots, x_d)$ ,  $i \in \{1, \dots, d\}$  and  $\alpha, \beta \in \mathcal{S}(d)$ , we denoted by  $[\mathbf{x}]_i$  the  $i^{\text{th}}$  component of  $\mathbf{x}$ , we defined  $\mathbf{x}_\alpha$  as the  $d$ -tuple  $(x_{\alpha_1}, \dots, x_{\alpha_d})$ , and we noticed that  $(\mathbf{x}_\alpha)_\beta = \mathbf{x}_{\alpha\beta}$ . (See Definition [Appendix A.2](#).)

**Fact Appendix B.2.** For any  $\mathbf{x} \in [n]^d$ ,  $\alpha \in \mathcal{S}(d)$ , and  $i, j \in \{1, \dots, d\}$ :

- (a)  $\mathbf{x}_{\alpha.ij} = (\mathbf{x}_\alpha)_{(ij)}$ .  
(b)  $[\mathbf{x}_{\alpha.ij}]_j = [\mathbf{x}]_d$ .

PROOF. (a) From  $\alpha.ij = \alpha(ij)$  and  $\mathbf{x}_{\alpha\beta} = (\mathbf{x}_\alpha)_\beta$ . (b) By induction on  $\alpha$ : if  $\alpha = \text{id}$ , then necessarily  $i = d$  and  $[\mathbf{x}_{\text{id}.dj}]_j = [\mathbf{x}]_d$  clearly holds. Otherwise,  $[\mathbf{x}_{\alpha.ij}]_j = [(\mathbf{x}_\alpha)_{ij}]_j = [\mathbf{x}_\alpha]_i = [\mathbf{x}]_d$ , by induction hypothesis.  $\square$

Given a  $d$ -tuple  $\mathbf{x} = (x_1, \dots, x_d)$  of first-order variables, we denote by  $\mathbf{x}^-$  the  $(d-1)$ -tuple obtained from  $\mathbf{x}$  by erasing its last component. That is,

$$(x_1, \dots, x_{d-1}, x_d)^- := (x_1, \dots, x_{d-1}).$$

In particular, for  $\alpha \in \mathcal{S}(d)$ , we denote by  $\mathbf{x}_\alpha^-$  the  $(d-1)$ -tuple  $(x_{\alpha_1}, \dots, x_{\alpha_{d-1}})$ . Each  $(d-1)$ -tuple build upon the  $d$  variables  $x_1, \dots, x_d$  can clearly be written  $\mathbf{x}_\alpha^-$  for some  $\alpha \in \mathcal{S}(d)$ . Therefore, the occurrence of a non-sorted atom in  $\Phi$  has the form  $Q(\mathbf{x}_\alpha^-)$  for some  $Q \in (Q_a)_{a \in \Sigma}$  and some  $\alpha \in \mathcal{S}(d)$ , and the purpose of this section amounts to rewrite each such occurrence  $Q(\mathbf{x}_\alpha^-)$  as  $Q'(\mathbf{x})$  for some well chosen relation  $Q'$ .

**Definition Appendix B.3.** Given a  $(d-1)$ -ary relation  $Q$  and two family of  $d$ -ary relations,  $(T_\alpha)_{\alpha \in \mathcal{S}(d)}$  and  $(Q_\alpha)_{\alpha \in \mathcal{S}(d)}$ , we say that  $(T_\alpha, Q_\alpha)_{\alpha \in \mathcal{S}(d)}$  is a  **$d$ -simulation of  $Q$**  if the following axioms hold, for any  $\alpha \in \mathcal{S}(d)$  and any  $i, j \leq d$  such that  $\alpha.ij$  is defined, and for any  $d$ -tuple  $\mathbf{x}$  of variables:

$$(A_1) \quad T_{id}(\mathbf{x}) = Q(\mathbf{x}^-) \text{ if } [\mathbf{x}]_d = 0.$$

$$(A_2) \quad T_{\alpha.ij}(\mathbf{x}) = T_{\alpha.ij}(\mathbf{x}_{(ij)}).$$

$$(A_3) \quad T_\alpha(\mathbf{x}) = T_{\alpha.ij}(\mathbf{x}) \text{ if } [\mathbf{x}]_i = 0.$$

$$(A_4) \quad Q_{id}(\mathbf{x}) = Q(\mathbf{x}^-).$$

$$(A_5) \quad Q_{\alpha.ij}(\mathbf{x}) = T_{\alpha.ij}(\mathbf{x}) \text{ if } [\mathbf{x}]_j = 0.$$

$$(A_6) \quad Q_{\alpha.ij}(\mathbf{x}) \text{ doesn't depend on } [\mathbf{x}]_j.$$

**Lemma Appendix B.4.** Let  $(T_\alpha, Q_\alpha)_{\alpha \in \mathcal{S}(d)}$  be a  $d$ -simulation of some  $(d-1)$ -ary predicate  $Q$ . For any  $\mathbf{x} \in [n]^d$  and  $\alpha \in \mathcal{S}(d)$ :  $Q_\alpha(\mathbf{x}_\alpha) = Q(\mathbf{x}^-)$ .

PROOF. Let us first prove that for any  $\mathbf{x} \in [n]^d$  and  $\alpha \in \mathcal{S}(d)$ :

$$[\mathbf{x}]_d = 0 \Rightarrow T_\alpha(\mathbf{x}_\alpha) = Q(\mathbf{x}^-). \quad (B.2)$$

We proceed by induction on  $\alpha$ . If  $\alpha = id$ , (B.2) follows from (A<sub>1</sub>). Given a non-identique permutation  $\alpha.ij$ , we have:

$$\begin{aligned} T_{\alpha.ij}(\mathbf{x}_{\alpha.ij}) &= T_{\alpha.ij}((\mathbf{x}_\alpha)_{(ij)}) \text{ by Fact Appendix B.2-(a).} \\ &= T_{\alpha.ij}(\mathbf{x}_\alpha) \text{ by (A}_2\text{).} \end{aligned}$$

But  $[\mathbf{x}_\alpha]_i = [(\mathbf{x}_\alpha)_{(ij)}]_j = [\mathbf{x}_{\alpha.ij}]_j$  and hence, by Fact Appendix B.2-(b):  $[\mathbf{x}_\alpha]_i = [\mathbf{x}]_d = 0$ . Therefore we can resume the above sequence of equalities with:

$$\begin{aligned} T_{\alpha.ij}(\mathbf{x}_{\alpha.ij}) &= T_\alpha(\mathbf{x}_\alpha) \text{ by (A}_3\text{) since } [\mathbf{x}_\alpha]_i = 0. \\ &= Q(\mathbf{x}^-) \text{ by induction hypothesis.} \end{aligned}$$

This completes the proof of (B.2)

Let us now prove the equality  $Q_\alpha(\mathbf{x}_\alpha) = Q(\mathbf{x}^-)$ . If  $\alpha = id$ , the result comes from (A<sub>4</sub>). For a non-identique permutation  $\alpha.ij$ , we have to prove  $Q_{\alpha.ij}(\mathbf{x}_{\alpha.ij}) = Q(\mathbf{x}^-)$  for any tuple  $\mathbf{x} \in [n]^d$ . First notice that we can restrict, without loss of generality, to the case where  $[\mathbf{x}]_d = 0$ . Indeed, we can otherwise consider the tuple  $\mathbf{y}$  obtained from  $\mathbf{x}$  by setting  $[\mathbf{x}]_d$  to 0. (That is,  $\mathbf{y}$  only differs from  $\mathbf{x}$  by its  $d^{th}$  component, which is null.) Clearly,  $\mathbf{y}^- = \mathbf{x}^-$ . Besides, the  $j^{th}$  component of  $\mathbf{x}_{\alpha.ij}$  is  $[\mathbf{x}]_d$ , from Fact Appendix B.2-(b). Similarly, the  $j^{th}$  component of  $\mathbf{y}_{\alpha.ij}$  is  $[\mathbf{y}]_d$ . Hence, the tuples  $\mathbf{x}_{\alpha.ij}$  and  $\mathbf{y}_{\alpha.ij}$  coincide on each component of rank distinct from  $j$ . Therefore  $Q_{\alpha.ij}(\mathbf{x}_{\alpha.ij}) = Q_{\alpha.ij}(\mathbf{y}_{\alpha.ij})$  by (A<sub>6</sub>) and we get:  $Q_{\alpha.ij}(\mathbf{x}_{\alpha.ij}) = Q(\mathbf{x}^-)$  iff  $Q_{\alpha.ij}(\mathbf{y}_{\alpha.ij}) = Q(\mathbf{y}^-)$ . Thus, we can assume  $[\mathbf{x}]_d = 0$ . It follows  $[\mathbf{x}_{\alpha.ij}]_j = [\mathbf{x}]_d = 0$ , by Fact Appendix B.2-(b), and hence:

$$\begin{aligned} Q_{\alpha.ij}(\mathbf{x}_{\alpha.ij}) &= T_{\alpha.ij}(\mathbf{x}_{\alpha.ij}) \text{ by (A}_5\text{), since } [\mathbf{x}_{\alpha.ij}]_j = 0. \\ &= Q(\mathbf{x}^-) \text{ by (B.2), since } [\mathbf{x}]_d = 0. \end{aligned}$$

The proof is complete.  $\square$



**Lemma Appendix B.5.** Let  $Q$  be a  $(d-1)$ -ary relation and  $(T_\alpha)_{\alpha \in S(d)}$ ,  $(Q_\alpha)_{\alpha \in S(d)}$  be two tuple of  $d$ -ary relations satisfying, for each  $d$ -tuple  $\mathbf{x} = (x_1, \dots, x_d)$ :

$$(F_1) \min(x_d) \rightarrow (T_{id}(\mathbf{x}) \leftrightarrow Q(\mathbf{x}^-)).$$

$$(F_2) (\neg \max(x_i) \wedge \neg \max(x_j)) \rightarrow (T_{\alpha.ij}(\mathbf{x}^{(i)}) \leftrightarrow T_{\alpha.ij}(\mathbf{x}^{(j)})).$$

$$(F_3) \min(x_i) \rightarrow (T_\alpha(\mathbf{x}) \leftrightarrow T_{\alpha.ij}(\mathbf{x})).$$

$$(F_4) Q_{id}(\mathbf{x}) \leftrightarrow Q(\mathbf{x}^-).$$

$$(F_5) \min(x_j) \rightarrow (Q_{\alpha.ij}(\mathbf{x}) \leftrightarrow T_{\alpha.ij}(\mathbf{x})).$$

$$(F_6) Q_{\alpha.ij}(\mathbf{x}) \leftrightarrow Q_{\alpha.ij}(\mathbf{x}^{(j)}).$$

Then  $(T_\alpha, Q_\alpha)_{\alpha \in S(d)}$  is a  $d$ -simulation of  $Q$ . Furthermore, each  $Q$  admits such a  $d$ -simulation fulfilling  $(F_1) \dots (F_6)$ .

**PROOF.** Clearly, the formula  $(F_i)$  is a mere transcription of the axiom  $(A_i)$  for each  $i \neq 2$ . We have to prove that  $(F_2)$  implies  $(A_2)$ . Formula  $(F_2)$  yields that  $T_{\alpha.ij}(\mathbf{x})$  has the same value on tuples of the form

$$\mathbf{x} = (\mathbf{u}, x+c, \mathbf{v}, y-c, \mathbf{w}),$$

for any  $c \leq \min\{n-1-x, y\}$ , where  $x+c$  (resp.  $y-c$ ) is the component of rank  $i$  (resp.  $j$ ) of  $\mathbf{x}$ . (That is:  $\mathbf{u} \in [n]^{i-1}$ ,  $\mathbf{v} \in [n]^{j-i-1}$  and  $\mathbf{w} \in [n]^{d-j}$ .) In other words, the value of  $T_{\alpha.ij}(\mathbf{x})$  depends on  $[\mathbf{x}]_i + [\mathbf{x}]_j$  rather than on the precise values of these two components. As a consequence, for any  $\mathbf{u} \in [n]^{i-1}$ ,  $\mathbf{v} \in [n]^{j-i-1}$ ,  $\mathbf{w} \in [n]^{d-j}$  and any  $x, y \in [n]$ :

$$T_{\alpha.ij}(\mathbf{u}, x, \mathbf{v}, y, \mathbf{w}) = T_{\alpha.ij}(\mathbf{u}, y, \mathbf{v}, x, \mathbf{w}).$$

This is Axiom  $(A_2)$ .

It remains to prove that such relations  $(T_\alpha)_{\alpha \in S(d)}$  and  $(Q_\alpha)_{\alpha \in S(d)}$  exist for every  $Q$ . To see this, assume we are given a  $(d-1)$ -ary relation  $Q$  and define the  $Q_\alpha$ 's and the  $T_\alpha$ 's as follows:

- $Q_\alpha(\mathbf{x}) = Q(\mathbf{x}_{\alpha-1}^-)$  for any  $\mathbf{x} \in [n]^d$ ;
- $T_{id}(\mathbf{x}) = Q_{id}(\mathbf{x})$  for any  $\mathbf{x} \in [n]^d$ ;
- $T_{\alpha.ij}(\mathbf{u}, x, \mathbf{v}, y, \mathbf{w}) = Q_{\alpha.ij}(\mathbf{u}, x+y, \mathbf{v}, 0, \mathbf{w})$  for any  $x, y \in [n]$ ,  $\mathbf{u} \in [n]^{i-1}$ ,  $\mathbf{v} \in [n]^{j-i-1}$  and  $\mathbf{w} \in [n]^{d-j}$ .

We leave it to the reader to check that the sequence  $(T_\alpha, Q_\alpha)_{\alpha \in S(d)}$  satisfies the formulas  $(F_1), \dots, (F_6)$  (and hence, is a  $d$ -simulation of  $Q$ ).  $\square$

**Proposition Appendix B.6.** For any  $d > 0$ ,  $\text{ESO}(\forall^d, \text{arity } d, \text{half-sorted}) \subseteq \text{ESO}(\forall^d, \text{arity } d, \text{sorted})$  on  $\text{COORD}^{d-1}$ .

Propositions [Appendix A.9](#) and [Appendix B.6](#) can now be collected in the following result:

**Theorem Appendix B.7.**  $\text{ESO}(\forall^d, \text{arity } d) = \text{ESO}(\forall^d, \text{arity } d, \text{sorted})$  on  $\text{COORD}^{d-1}$  for any  $d > 1$ .